Cambridge Tracts in Mathematics
and Mathematical Physics

General Editors
G. H. HARDY, M.A., F.R.S.
E. CUNNINGHAM, M.A.

No. 21

THE THEORY OF INTEGRATION

CAMBRIDGE
UNIVERSITY PRESS
LONDON : Fetter Lane

NEW YORK
The Macmillan Co.
BOMBAY, CALCUTTA and
MADRAS
Macmillan and Co., Ltd.
TORONTO
The Macmillan Co. of
Canada, Ltd.
TOKYO
Maruzen-Kabushiki-Kaisha

THE THEORY OF INTEGRATION

BY

L. C. YOUNG

CAMBRIDGE
AT THE UNIVERSITY PRESS
1927

PRINTED IN GREAT BRITAIN

PREFACE

THE object of small treatises of this kind is to enable the general student to gain rapid access to the various branches of Modern Mathematics, thereby preventing this science from breaking up into a number of disconnected parts, each belonging to its own specialist and closed to the outsider. Mathematics must form a whole, any progress in one of its parts stimulating advance in the others and raising new problems; when a branch is severed from the tree, it dies.

In writing this book, I have therefore tried above all to simplify the work of the student. On the one hand, practically no knowledge is assumed (merely what concerns existence of real numbers and their symbolism); on the other hand, the ideas of Cauchy, Riemann, Darboux, Weierstrass, familiar to the reader who is acquainted with the elementary theory, are used as much as possible.

I have also hopes that it will be of some use to the initiated, who may find here new points of view and greater generality in the treatment, owing to the idea of integration with respect to a function in space of n-dimensions. I have not however included what Hobson and others call the Fundamental Theorem of the Integral Calculus, namely the connection with the Theory of Derivation.

The Theory of Integration, which forms the subject of this book, has long been one of the most useful tools of Mathematics. Its methods were already employed with success by the ancient Greeks, in their investigations about Areas and Volumes. They possessed the method of exhaustion, the method of series. They were very clear about the idea of limit and this perhaps made them suspicious of the unsound method of infinitesimals, as results thus obtained were always established independently.

After the Dark Ages the rediscovery of this last method and the use of the symbolism of Algebra rendered possible the creation of the Calculus by Newton and Leibnitz. Unfortunately, *believing they had reduced everything to symbols*, they did not realise the need of examining the *ideas* these represented and testing their soundness. They conceived their Calculus to be purely formal; limiting process, a mere operation on symbols. They were very clear as to the properties they expected of such operations: possibility of operation, existence of inverse operation,

reversibility of order of two consecutive operations. But they were not so clear as to the properties implied of the entities operated on, which in their case were *functions*, loosely defined as numbers depending on variable quantities.

This conception of Mathematics persisted for a long time. In the nineteenth century, however, the feeling that Mathematics is not the entire property of the mere Calculator, but rather that of the Thinker, revived at last. The result was the development of Geometry, the Theory of Groups, the Theory of Vectors, the systematic use of the Imaginary.

The Mathematicians of that century naturally also perceived the need of reforming the Infinitesimal Calculus. The reform was started with Cauchy's Theory of Limits, based on Inequalities. Cauchy also introduced the notion of *Continuity* and attempted to use it as a foundation for the Calculus. He saw the unsatisfactoriness of the notion, hitherto adopted, of Integration as Inverse Differentiation: a definition which is not constructive requires a theorem ensuring the existence and unicity of the entity in question. Cauchy defined Integration for a continuous function by an always possible limiting process and he proved that it could be considered as the inverse of Differentiation.

But the occurrence of discontinuous functions in certain simple problems and the discovery, by Weierstrass, of continuous non-differentiable functions,—by Riemann, of discontinuous integrable functions, showed that continuity is both inconvenient and unnatural for the foundation of the Calculus.

The Theory of Integration and that of Differentiation have since been built up separately as parts of the New Calculus, the *Calculus of Real Functions*, whose great generality, far from being due to a love of complication on the part of its founders, as was at one time asserted, is to be attributed to the simplicity and straightforwardness of its methods

This New Calculus would never have been possible but for the wonderful ideas of Cantor, at first completely unintelligible to the Mathematicians of the Older School, some of whom even wilfully misunderstood them and sought to lead others into error with regard to them; but which, fortunately, very much influenced a few younger men since become famous. Only too often have ideas of the greatest value been left for a long time unheeded, while their authors remain in obscurity. Galois, the founder of the Theory of Groups, was ploughed at the entrance examination of the École Polytechnique through knowing more than his examiners, and this was only the first of a series of disappointments which embittered his short life. Grassmann, the creator of the

theory of vectors, remained a schoolmaster most of his life, and his book, the "Ausdehnungslehre," was burned by the publishers, who could find no buyers.

Cantor's Theory of Sets of Points and Borel's improvement of the theory of content or measure of such sets paved the way to the semi-geometric definitions of Integration, given almost simultaneously by Lebesgue and Young. These definitions represent an extension comparable to that of Arithmetic on the introduction of irrational numbers. They are substantially equivalent to the more direct one here adopted, later given by Young, using the work of Baire on functions.

My father had long thought of writing a connected account of his theory. In carrying out this task myself at his suggestion, I have tried to do justice to his ideas and to introduce a few minor improvements of my own. If I have succeeded in my endeavours, it will have been largely owing to his encouragement, and to the constant assistance of my mother and of my sister Miss R. C. H. Young.

L. C. Y.

September 1926.

CONTENTS

PAGE

PREFACE v

CHAP.

I. THE METHOD OF MONOTONE SEQUENCES

§ 1. Successions of numbers 1

§ 2. Successions of functions 3

§ 3. Limits of a function at a point 5

§ 4. Semicontinuity and the theorem of bounds . . 7

II. THE GENERATION OF FUNCTIONS

§ 1. The simple functions 9

§ 2. The generation of general semicontinuous functions by monotone sequences of simple functions 10

§ 3. The generation of new functions 12

§ 4. Functions defined by other methods 14

III. MONOTONE FUNCTIONS

§ 1. Monotone functions of one variable 16

§ 2. Monotone functions of two variables 16

§ 3. Monotone functions of n variables 18

§ 4. Total increment of a monotone function over a stretch . 19

§ 5. The integration of functions constant in stretches with respect to a monotone increasing function 20

IV. THE INTEGRATION OF FUNCTIONS

§ 1. Methods of the theory of integration 22

§ 2. Darboux's theorem 24

§ 3. The integration of semicontinuous functions . . . 27

V. THE INTEGRATION OF FUNCTIONS (*continued*)

§ 1. Final extension of integration 29

§ 2. Indefinite, parametric, and repeated integration . . 33

§ 3. Term-by-term integration of sequences and successions . 35

§ 4. Theorems concerning the integrator . . . 39

APPENDIX

1. Continuity and uniform convergence 42
2. The generation of sets of points, and the theory of content . 46
3. Functions defined in a set of points 48

INDEX 53

CHAPTER I

THE METHOD OF MONOTONE SEQUENCES

§ 1. Successions of numbers*.

A set of numbers is said to be *bounded* if there exist a finite number greater and a finite number less than all those of the set. The smallest number such that no number greater than it belongs to the set is called the *upper bound* of the set of numbers; the greatest number such that no smaller number belongs to the set is called the *lower bound* of the set. If there is no finite number greater than all those of the set, the set of numbers is said to be *unbounded above*, and we shall agree to say that its upper bound is $+\infty$; similarly, if there is no finite number less than all those of the set, the set of numbers is said to be *unbounded below*, and its lower bound will be $-\infty$.

A countably infinite set of numbers, written down in a definite order, is called a *succession*, e.g.

$$a_1, a_2, \ldots, a_n, \ldots .$$

Repetition of a number is allowed.

A succession is said to be *monotone ascending* if each term is greater than or equal to the preceding,

$$a_1 \leqslant a_2 \leqslant \ldots \leqslant a_n \leqslant \ldots;$$

a succession is said to be *monotone descending* if each term is less than or equal to the preceding,

$$a_1 \geqslant a_2 \geqslant \ldots \geqslant a_n \geqslant \ldots .$$

In either case the succession is said to be monotone, or to be a *monotone sequence*. One of the bounds of a monotone sequence is clearly its first term a_1. The other is called the *unique limit* of the monotone sequence, whether finite or infinite.

Given any succession S

$$a_1, a_2, \ldots, a_n, \ldots,$$

the succession S_r obtained from it by leaving out its r first terms, has its bounds K_r, k_r lying between those of the succession S_{r-1}, i.e. we have

$$k_{r-1} \leqslant k_r \leqslant K_r \leqslant K_{r-1}.$$

Thus, the succession of the upper bounds

$$K_1, K_2, \ldots, K_n, \ldots$$

* We are dealing with real numbers; $+\infty$ and $-\infty$ are regarded as distinct.

is monotone descending. Its unique limit is called the *upper limit* of the succession S. Similarly, the unique limit of the monotone ascending sequence of the lower bounds is called the *lower limit* of S.

The upper and lower limits of a monotone sequence are equal and coincide with its unique limit.

For one succession of bounds coincides with the given sequence and the other consists of terms all equal to the unique limit.

A succession whose upper and lower limits are equal is called a *sequence*, and their common value is called the *unique limit* of the sequence. A sequence having a finite limit is said to converge; otherwise it diverges. A succession which is not a sequence is said to oscillate.

Given any succession S, we call subsuccession of S a succession of numbers all belonging to S and occurring in the same order as in S. We call subsequence of S any subsuccession of S which is a sequence. The unique limit of any subsequence of S is called *a limit* of S.

THEOREM. The upper and lower limits of a succession are limits of that succession.

It is sufficient to prove this for the upper limit.

If the numbers K_n all belong to the succession, the theorem is obvious.

If any one of them does not belong to the succession, all the following are equal to it, and it is the upper limit*. Let r_0 be its index, r_1 that of the first term after a_{r_0} which is greater than a_{r_0}, r_2 that of the first term after a_{r_1} which is greater than a_{r_1}, and so on. Then the subsuccession of S,

$$a_{r_0}, a_{r_1}, \dots,$$

is a monotone sequence whose unique limit is K_{r_0}, because, by construction, no term of S after a_{r_0} can exceed this limit which, being the upper bound of a subsequence of S_{r_0}, cannot exceed K_{r_0}.

THEOREM. The upper and lower limits of a subsuccession of S lie between those of S.

Let S' be the given subsuccession. Let S_r' be the succession obtained from S' by leaving out its r first terms, and let S_r be obtained similarly from S.

Then S_r' is a subsuccession of S_r, and its bounds K_r' and k_r' lie between those of S_r, K_r, and k_r,

$$k_r \leqslant k_r' \leqslant K_r' \leqslant K_r.$$

This holds for all r; the theorem follows.

* Omitting from a set of numbers a finite number of its elements none of which coincide with the upper bound, does not affect the upper bound.

COROLLARY. All the limits of a succession lie between its upper and its lower limits*.

As a kind of converse of the preceding theorem, we have:

THEOREM. If a finite number of subsuccessions of S together contain all the elements of S, they have among their upper and lower limits those of S.

It is sufficient to prove this for the case of two subsuccessions and to consider only the upper limits. Let U be the upper limit of S. Then there is a subsequence having U as limit and we can so arrange that it belongs entirely to one of our subsuccessions (see footnote*). U is therefore a limit of that subsuccession. By the preceding theorem and its corollary it must therefore be its upper limit.

e-definitions of limits and convergence. To define by this method the upper limit of a bounded succession S,

$$a_1, a_2, \ldots, a_n, \ldots,$$

U is said to be the upper limit if given any positive number e however small, an index N can be found such that from and after $n = N$,

$$a_n \leqslant U + e,$$

while an infinite succession of n's can be found such that

$$a_n \geqslant U - e.$$

Similar definition for lower limit.

Thence the e-definition of limit of a convergent sequence

$$U - e \leqslant a_n \leqslant U + e,$$

for all n from and after N.

This method, which was known to the Greeks, is probably familiar to the student, who will easily prove the equivalence of the definitions so obtained with our former ones.

The characteristic advantage of our method is *to reduce the consideration of all successions to that of monotone sequences.*

§ 2. Successions of functions.

Corresponding to any set of numbers, we have on the straight line a set of points†; we need only choose an origin, a sense, and a unit of length.

We shall say a point is a *limiting point* of our set of points, and that the corresponding number is a limit of the set of numbers, if every interval of which it is the centre contains an infinite number of points of the set. This agrees with our definition of limit in the case of a succession.

* In particular, all the limits of a sequence coincide with its unique limit.

† This is equivalent to what is called the Cantor-Dedekind axiom.

A set containing all its limiting points is said to be *closed*. A point belonging to a set and not a limiting point is called an *isolated point* of that set.

After the finite sets, consisting of a finite number of points, and the sequences and successions of points, the simplest sets are the *intervals*. An interval consists of all the points between its endpoints. If it includes these it is closed, if neither, *open*.

Corresponding to any pair of numbers, we have a point in the plane; corresponding to any set of pairs of numbers, a set of points in the plane. Similarly any set of numbers given n by n, may be taken to represent a set of points *in n dimensions*. We may agree to represent the n coordinates of a point by *a single symbol*, and let x stand for the ensemble of the n numbers x_1, x_2, ..., x_n.

Two points $a=(a_1, a_2, \ldots, a_n)$ and $b=(b_1, \ldots, b_n)$ such that each coordinate of a is less than the corresponding one of b, define an interval (a, b), consisting of all the points x whose coordinates all lie between the corresponding coordinates of a and b, that is the set of points belonging to the rectangle or to the n-dimensional parallelopiped whose sides are parallel to the axes and which has a and b for its endpoints. The interval is open if it consists only of the internal points, closed, if of all the points. We shall call length of an interval the length of a principal diagonal. It is convenient to define a *stretch* to be either an *open* interval or the limiting case of an open interval when one or more of the inequalities between the coordinates of a and b are replaced by equalities. Thus in one dimension a stretch would be either an open interval or a point; in two dimensions it would be an open interval, or an open side of an interval, or a point, etc.

When distinction is necessary, we use the symbol t to denote a point on the straight line.

Definition of a function. A quantity y is said to be a function of x, in an interval (a, b), if to each x of that interval corresponds a single value y.

We use the symbol $f(x)$ to denote a function of x.

A function is said to be *bounded* if the set of numbers consisting of all its values is bounded.

If corresponding to each point x we are given a succession of numbers $f_1(x)$, $f_2(x)$, ..., these may be said to define a *succession of functions*.

The upper bound $K(x)$ of the numbers corresponding to the point x define a new function, which we call the *upper bounding function* of the succession of functions.

Similarly we define the *lower bounding function*.

The upper limit $U(x)$ of the numbers $f_1(x), f_2(x)$, ... corresponding to the point x defines a new function, which we call the *upper function* of the succession of functions.

Similarly we define the *lower function*.

The succession of functions is called a *sequence*, if its upper and lower

functions are identical. They are then called the *limiting function* of the sequence.

The succession of functions is said to be a *monotone ascending sequence* if the numbers corresponding to each x form an ascending sequence; it is said to be *monotone descending* if they form a monotone descending sequence. In either case the succession of functions is called a *monotone sequence of functions.*

Given any succession of functions, the upper bounding function $K_r(x)$ of the succession obtained by leaving out the r first functions, generates as r increases a monotone descending sequence which converges to the upper function $U(x)$, and the corresponding monotone ascending sequence of the lower bounding functions, converges to the lower function.

We have thus again reduced the consideration of all successions of functions to that of monotone sequences of functions.

In the case of monotone sequences of functions one bounding function is the first function, the other is the limiting function.

The convergence of a sequence of functions is said to be *bounded* if the bounding functions are bounded. The functions of the sequence are then said to be uniformly bounded.

The convergence of a sequence of functions $f_n(x)$ to a limiting function $f(x)$ in a closed interval is said to be *uniform* if, given any positive number e, an index N independent of x can be found such that, for all n from and after N,

$$|f_n(x) - f(x)| < e \text{ for all } x.$$

It is uniform in an open interval if uniform in every closed component interval. (See Appendix, 1, p. 44, last 18 lines and seq.)

§ 3. Limits of a function at a point. The upper and lower bounds of a set of numbers consisting of all the values of a function in an interval are called *the upper and lower bounds of the function in that interval.* Between them lie the bounds of the function in any interval interior to that interval.

Let x_0 be any point interior to the interval of definition (a, b). Let $w_1, w_2, \dots, w_r, \dots$ be any succession of intervals having x_0 as internal point and whose lengths converge to zero, each interval being contained in the preceding. Let M_r, m_r denote respectively the upper and lower bounds of the values of the function in w_r, *excluding the point* x_0. As r increases they form two monotone sequences of numbers and their limits, U and L respectively, are called the *upper* and the *lower limits*

of the function at the point x_0. These limits are independent of the choice of the succession of intervals*.

For consider two successions of intervals $w_1, w_2, \ldots$ and $w_1', w_2', \ldots$; let the corresponding limits be U, L and U', L'.

There cannot be more than a finite number of intervals of the second succession not interior to w_r. Since r is arbitrary U', L' therefore lie between U and L.

But, reversing the rôles of the two successions, U, L lie between U', L'. Hence $U = U'$, $L = L'$.

If in the definition of upper and lower limits at a point x_0 instead of intervals containing the point x_0, we consider intervals (open or closed) having x_0 for a corner-point and as before each contained in the preceding and with length decreasing to zero—this defines corresponding to each quadrant† (open or closed) at x_0 a unique pair of *upper* and *lower limits* (or limits of approach) *in that* (open or closed) *quadrant.*

In the case of an internal point of the interval of definition the upper and lower limits at the point are respectively equal to the greatest and to the least of the upper and lower limits in the various *closed* quadrants at the point. We may use this to define the upper and lower limits at every point of the *closed* interval of definition.

We thus have, in all cases, corresponding to each point x of the interval of definition, three numbers, the value of the function and its upper and lower limits.

It is obvious that the greatest and the least‡ of these three numbers are respectively the limits of the upper and lower bounds of the function in any succession of intervals of lengths decreasing to zero, the point x being internal to all of them (the value at x is not excluded). If the

* Also each interval need not lie in the preceding provided they ultimately shrink up to x_0. For let u_1 be the smallest interval containing all those of the succession, u_2 the smallest interval containing all except w_1, etc. Then $u_1, u_2, \ldots$ are each inside the preceding and shrink up to x_0. Again if v_1 is w_1, v_2 the largest interval inside w_2 which has no points outside $w_1, \ldots, v_n$ the largest interval in w_n having no points outside v_{n-1}, then $v_1, v_2, \ldots$ are a succession of intervals each inside the preceding and shrinking up to x_0. Also clearly

$$u_n \geqslant w_n \geqslant v_n.$$

Hence the limits for w_n lie between the corresponding limits for u_n and v_n which coincide.

† By quadrant at a point we mean, in one dimension, one or other of the two sides of the point; in n dimensions, an angle determined by n parallels to the axes through the point.

‡ They are sometimes called the maximum and minimum limits and their difference is sometimes called the *jump* of the function at the point x.

value at the point is the greatest, it is also the limit of the upper bounds of the function in any succession of closed intervals which have x as corner-point, and whose lengths decrease to zero. In that case the function is said to be *upper semicontinuous* at the point x. Similarly, if the value of the function is the least of the three numbers, the function is said to be *lower semicontinuous* at the point.

Similarly we may define upper and lower semicontinuity in a closed or open quadrant at a point.

If a function is both upper and lower semicontinuous at a point and its value there is finite, it is said to be *continuous** at the point. Otherwise the function is *discontinuous* at the point.

§4. **Semicontinuity and the theorem of bounds.** A function is said to be upper semicontinuous in an interval if it is upper semicontinuous at every point of the interval. We shall call it a *U-function.*

A function is said to be lower semicontinuous in an interval if it is lower semicontinuous at every point of the interval. We shall call it an *L-function.*

In either case it is said to be *semicontinuous* in the interval. A function which is both an L and a U and assumes only finite values is said to be *continuous.*

e-definition of semicontinuity at a point. A function $f(x)$ is said to be upper semicontinuous at the point x_0 if, given any positive quantity e, there is an interval having x_0 as middle point throughout which if $f(x_0)$ is finite

$$f(x) \leqslant f(x_0) + e,$$

while, if $f(x_0)$ is $-\infty$,

$$f(x) \leqslant -1/e \qquad (U).$$

A function $f(x)$ is said to be lower semicontinuous at the point x_0 if, given any positive quantity e, there is an interval having x_0 as middle point, throughout which, if $f(x_0)$ is finite,

$$f(x) \geqslant f(x_0) - e,$$

while, if $f(x_0)$ is $+\infty$,

$$f(x) \geqslant 1/e \qquad (L).$$

If $f(x_0)$ is $+\infty$, then $f(x)$ is certainly upper semicontinuous at x_0; if $f(x_0)$ is $-\infty$, then $f(x)$ is certainly lower semicontinuous at x_0.

THEOREM. An L-function assumes its lower bound in every closed interval; a U-function, its upper bound.

Divide the given interval into two equal parts. If m is the lower bound of our L-function in the given closed interval, it is also its lower

* See Appendix, 1.

bound in at least one of these closed subintervals. Let W_1 be the first having this property. Again divide W_1 into two equal parts, and let W_2 be the first of these closed subintervals of W_1 in which m is again the lower bound. And so on.

The succession of intervals

$$W_1,\ W_2,\ \ldots,\ W_n,\ \ldots$$

consists of closed intervals, each contained in the preceding, and whose lengths decrease to zero. There is exactly one point common to all of them. The value of our L-function at that point is therefore the limit of its lower bounds in the succession of intervals, that is to say m. Q.E.D.

Similarly, we may prove the corresponding theorem for U-functions.

The theorem of bounds.

If $f_1(x)$, $f_2(x)$, ..., is a monotone ascending sequence of functions having $f(x)$ as limiting function; if u_n and l_n are respectively the upper and the lower bounds of $f_n(x)$ in a fixed closed interval, u and l those of $f(x)$ in the same interval, then

$$\lim u_n = u \quad \text{while} \quad \lim l_n \leqslant l.$$

Moreover, if the functions are all L-functions, then

$$\lim l_n = l.$$

It is obvious that if f_1 is never greater than f_2, the same is true of their bounds. Hence

$$u_1 \leqslant u_2 \leqslant \ldots \leqslant u; \quad l_1 \leqslant l_2 \leqslant \ldots \leqslant l.$$

Therefore $$\lim u_n \leqslant u; \quad \lim l_n \leqslant l.$$

But if A is any quantity less than u, there are points x such that $f(x) > A$. At any such point x, we have, since $f_n(x)$ converges to $f(x)$, $f_n(x) > A$ from and after a certain $f_N(x)$. Therefore, a fortiori,

$$\lim u_n > A,$$

or, since A was any quantity less than u, $\lim u_n \geqslant u$.

Therefore $$\lim u_n = u, \quad \text{while} \quad \lim l_n \leqslant l.$$

In the case in which the functions f_n are L-functions, we can find a point x_n where f_n assumes its lower bound. Let x' be any limiting point of the x_n, and let B be any quantity less than $f(x')$. Then

$$f_n(x') > B$$

from and after a certain $f_{N'}$. There is therefore an interval surrounding x' throughout which, since $f_{N'}$ is an L-function, $f_{N'}(x)$ is greater than B. In that same interval, by monotony,

$$f_n(x) > B$$

from and after $f_{N'}$. In this interval there are an infinite number of the points x_n with indices greater than N'. Therefore, if N'' is the first of these, then

$$f_{N''}(x_{N''}) = l_{N''} > B.$$

Therefore, a fortiori,

$$\lim l_n \geqslant B.$$

Since B was any number less than $f(x')$,

$$f(x') \leqslant \lim l_n,$$

and, a fortiori,

$$l \leqslant \lim l_n.$$

CHAPTER II

THE GENERATION OF FUNCTIONS

§ 1. **The simple functions.** The simplest functions are the constants; the value of y is the same for all x.

The next simplest functions are *the functions constant in stretches*, whose interval of definition is the sum of a finite number of stretches inside each of which the function is a finite constant. By a stretch we mean, as explained on p. 4, an open interval or a kind of limiting case of an open interval, such as a point.

A function constant in stretches is not in general semicontinuous.

For example, the function defined in the interval (0, 1), whose value in the open interval (0, 1/2) is 0, whose value at the point 1/2 is 1/2, whose value in the open interval (1/2, 1) is 1, is not semicontinuous at the point 1/2.

A function constant in stretches will certainly be lower semicontinuous at every point of its interval of definition if its value in every stretch which is not an interval is equal to the least of the values in the neighbouring intervals. It is then called a *simple L-function.*

Similarly, we call *simple U-function* a function constant in stretches whose value in every stretch which is not an interval is the greatest of the values in the neighbouring intervals.

These two types of *simple functions* have the following properties:

(i) The sum of two functions of the same type is of that type.

(ii) The function equal to the greater of two functions of the same type, and the function equal to the smaller of two functions of the same type, at each point, are of that type.

(iii) Change of sign transfers each type to the twin type.

These properties we shall refer to as *the three fundamental properties of class.*

§ 2. The generation of general semicontinuous functions by monotone sequences of simple functions.

THEOREM. Any L-function bounded below is expressible as the limit of a monotone ascending sequence of simple L-functions and also as the limit of such a sequence of simple U-functions.

Let $f(x)$ be the given L-function, defined in the interval (a, b). Let m be its lower bound.

We divide (a, b) into two equal parts and we call m'_1, m''_1 the lower bounds of $f(x)$ in each of these parts, the common boundary points being taken to belong to both.

We again bisect each of the parts.

(In the case of several variables we bisect in turn the range of each of them.)

Let m'_2, m''_2, m'''_2, m''''_2, be the bounds so obtained, and so on.

If at any stage, the nth say, one of these numbers be infinite, we replace it by the greatest of all the preceding plus n. In that case $f(x)$ would have to be infinite ($+\infty$) in the whole of the corresponding interval.

Let $a_0(x)$, $b_0(x)$ be the constant m.

Let $a_1(x)$, $b_1(x)$ be respectively the simple L-function and the simple U-function which are equal to m'_1 in the first half of (a, b) and to m''_1 in the second half, their values at the common boundary points* being of course respectively the smallest and the largest of these two numbers.

Let $a_2(x)$, $b_2(x)$ be respectively the simple L-function and the simple U-function equal to m'_2 in the first quarter of (a, b), to m''_2 in the second, and so on.

The two sequences of functions,

$$a_0(x),\quad a_1(x),\quad a_2(x), \ldots,$$
$$b_0(x),\quad b_1(x),\quad b_2(x), \ldots,$$

are monotone ascending. The values of their limiting function at any point x are in both cases equal to the limits of the lower bounds of $f(x)$ in one or more successions of closed intervals each contained in the preceding and whose lengths tend to zero, and such that x is common to all of them.

Since $f(x)$ is semicontinuous the limiting functions therefore both coincide with $f(x)$. Q.E.D.

Similarly we can establish the corresponding

* At which alone they differ.

THEOREM. Any U-function bounded above is expressible as the limit of a monotone descending sequence of simple U-functions, and also as the limit of such a sequence of simple L-functions.

The converse of these theorems is not true. Monotone sequences of simple functions do not in general define semicontinuous functions. It is, however, easy to show that monotone ascending sequences of simple L-functions always define L-functions, while monotone descending sequences of simple U-functions always define U-functions. More generally, we have

THEOREM. The limiting function of a monotone ascending sequence of L-functions is an L-function.

Let $f_n(x)$ be the generic function of the sequence, and $f(x)$ the limiting function. Since f is never less than f_n, the same is true of their lower bounds in any interval and consequently of their lower limits at any point. Let l be the lower limit of $f(x_0)$ at the point x_0. Since f_n is semicontinuous, $f_n(x_0)$ is smaller than or equal to the lower limit of f_n at x_0. Therefore, a fortiori,

$$f_n(x_0) \leqslant l.$$

This is true for all n. Therefore

$$f(x_0) \leqslant l.$$

Similarly we can prove the corresponding

THEOREM. The limit of a monotone descending sequence of U-functions is itself a U-function.

The two types of semicontinuous functions are easily seen to possess the *three fundamental properties of class*.

i. The sum of two functions of the same type is of that type.

If f and g are two L-functions bounded below they are the limits of two monotone sequences of simple L-functions

$$f_1, f_2, \ldots, \quad g_1, g_2, \ldots,$$

their sum is therefore the limit of the monotone ascending sequence

$$f_1+g_1, \quad f_2+g_2, \ldots$$

of simple L-functions, i.e. an L-function*.

If f and g are not bounded below, every point, where neither assumes the value $-\infty$, is internal to an interval where both are bounded below.

At a point where one assumes this value, the other being different from $+\infty$, the sum assumes the value $-\infty$, and is consequently lower semicontinuous.

At a point where one function assumes the value $-\infty$, and the other the value $+\infty$, the sum is not defined.

Similarly we prove the theorem for U-functions.

* The same reasoning establishes the property for any type of functions defined by monotone sequences of functions belonging to a type which has the property.

ii. The function which is the greater of two functions of the same type is of that type; so is the function which is the lesser of two functions of the same type.

Let f and g be two L-functions bounded below, h the function which is equal to the greater (or to the smaller) of the two at each point. f and g are the limits of monotone ascending sequences of simple L-functions

$$f_1, f_2, \ldots, f_n, \ldots,$$

$$g_1, g_2, \ldots, g_n, \ldots.$$

At each point, h is the limit of the monotone ascending sequence of simple L-functions

$$h_1, h_2, \ldots, h_n, \ldots,$$

each of which is equal to the greater (or the smaller) of the two corresponding functions at each point. h is therefore an L-function.

Now let f and g be unbounded below.

Every point where neither is equal to $-\infty$ is internal to an interval in which the above holds.

A point where f is not $-\infty$, is internal to an interval where f is greater than a finite number m. The greater of f and g is not affected if we replace by m all the values of g in that open interval which are less than m; neither does this affect the lower semicontinuity of g.

In the remaining cases h is obviously lower semicontinuous.

Similarly the theorem is proved for U-functions.

iii. Change of sign transfers a type to the twin type.

If f is the limit of a monotone ascending sequence of simple L-functions f_n then $-f$ is the limit of the monotone sequence of simple U-functions $-f_n$.

At a point where f is $-\infty$, $-f$ is $+\infty$.

These properties are also easy to establish independently. The method adopted has the advantage of being applicable to a number of other theorems.

We have seen that all bounded semicontinuous functions are expressible as the limits of monotone sequences of simple functions.

We shall next show that all functions it has been found possible to define up to this date are expressible as the limits of a *system* of monotone sequences of these functions.

§ 3. **The generation of new functions.** Monotone sequences of semicontinuous functions fall into four classes. Two of these have been already seen to define semicontinuous functions, namely ascending sequences of L- and descending sequences of U-functions.

The remaining two classes usually define functions which are not semicontinuous.

DEFINITION. The limit of a monotone descending sequence of L-functions is called an *UL-function*. That of a monotone ascending sequence of U-functions is called a *LU-function*.

The prefix U to the name of a function denotes the limit of a *descending* sequence of functions of that name. The prefix L to the name of a function denotes the limit of an *ascending* sequence of functions of that name. According to this nomenclature,

$$LL = L, \quad UU = U.$$

THEOREM. An L-function is both a LU and an UL. So is a U-function.

We know already that an L is a LU (p. 10). It is also an UL, because a sequence of functions each equal to the following is a particular case of a monotone descending sequence.

There are functions which are both LU and UL without being either L or U, e.g. the function constant in stretches of p. 9. There are also functions which are LU's without being UL's.

The two types of functions possess the *three fundamental properties of class*. (See footnote, p. 11.)

Monotone sequences of functions of one of these types again fall into four classes. Their limiting functions are, according to our nomenclature, LLU, ULU, LUL, UUL. It is easy to show that only two of these are new types of functions.

Let $f_1 \leqslant f_2 \leqslant \ldots \to f$ be any monotone ascending sequence of LU-functions, and f their limiting function.

f_1 is the limit of an ascending sequence

$$f_{11} \leqslant f_{12} \leqslant \ldots \to f_1$$

of U-functions. f_2 is the limit of an ascending sequence

$$f'_{21} \leqslant f'_{22} \leqslant \ldots \to f_2$$

of U-functions. Since $f_2 \geqslant f_1$, it is also the limit of the monotone ascending sequence

$$f_{21} \leqslant f_{22} \leqslant f_{23} \leqslant \ldots \to f_2$$

of U functions such that $f_{2,r}$ is the greater of $f_{1,r}$ and $f'_{2,r}$ at every point. Similarly f_3 is the limit of a monotone ascending sequence

$$f_{31} \leqslant f_{32} \leqslant \ldots \to f_3$$

of U-functions, such that $f_{3,r} \geqslant f_{2,r}$.

Proceeding in this way, we get

$$\begin{array}{ll} f_{11} \leqslant f_{12} \leqslant \ldots \to & f_1 \\ f_{21} \leqslant f_{22} \leqslant \ldots \to & f_2 \\ \ldots\ldots\ldots\ldots\ldots & \downarrow \\ \ldots\ldots\ldots\ldots\ldots & f \end{array}$$

The monotone ascending sequence

$$f_{11} \leqslant f_{22} \leqslant \dots$$

has all its terms, and therefore its limit also, less than or equal to f. On the other hand, each of its terms from the rth onwards is greater than or equal to the corresponding term of the sequence which has for limit f_r. It follows that

$$f_{11} \leqslant f_{22} \leqslant \dots \to f,$$

and since the functions are all U-functions, this proves that f is a LU-function, i.e. $LLU = LU$. And similarly $UUL = UL$.

THEOREM. A LU-function is both an ULU and a LUL. So is an UL.

A LU is obviously an ULU. It is a LUL being an L-U, i.e. an L-UL, since a U is an UL.

The *three fundamental properties of class* hold for the two types.

It is clear that we can go on defining functions in this way *ad infinitum*. At each stage we have four kinds of monotone sequences, and only two of these define new types of functions. Each type of function is a particular case of both types of the next stage. At each stage the three fundamental properties of class hold for the two types.

For the sake of brevity, we call U_n and L_n the two types of the nth stage, n being the number of sequences necessary to generate either, starting with simple functions.

Two new types of functions may then be obtained as follows. Let $f_1, f_2, \dots$ be a monotone sequence such that f_1 belongs to one of the types of the first stage, f_2 to one of the second stage, and so on. Its limiting function will in general belong to none of the previous types. The two types thus defined we call U_ω and L_ω; monotone sequences of these give again two more general types. And so on.

§ **4. Functions defined by other methods.** For defining a function in a given interval, the only means at our disposal are the operations of arithmetic and that of passage to the limit applied to known functions, such as the variables x and constants. This may be further complicated by choosing different laws in different parts of the fundamental interval. As we are using the operation of passage to the limit, we may always suppose these parts to consist of a finite number of stretches.

The operations of arithmetic and of passage to the limit are called *analytic operations*. A function is said to be *representable analytically*, or to be capable of an *analytic expression*, if its value at each point of

its interval of definition may be obtained by means of the same analytic operations on the variables x and on constants. When considering analytic operations, we can leave out those of subtraction and division, as the former may be replaced by multiplication by a constant (-1) and addition, the latter by a power series.

THEOREM. Analytic operations performed on functions defined by a system of monotone sequences of simple functions lead to functions of the same kind.

(*a*) The operation of passage to the limit merely involves expressing the function as the limit of a sequence of functions already defined. Let these be

$$f_1, \quad f_2, \ldots.$$

The monotone ascending sequence

$$g_1, \quad g_2, \ldots,$$

where $g_1 = f_1$, and g_n is the greater of f_n and g_{n-1}, by property ii consists of functions of the same type as the former sequence. Its limit is the upper bounding function of that sequence, and belongs at most to the next stage. The same is true of the upper bounding functions of the sequences obtained by leaving out the r first functions f. These upper bounding functions form a monotone sequence whose limit is the required function. This proves our theorem for the operation of passage to the limit.

(*b*) If f is a function obtainable by monotone sequences, the same is true of its product by a function constant in stretches. Its product by any other function obtainable by monotone sequences may then be obtained by operations of passage to the limit.

(*c*) The operation of addition has already been seen not to affect a type. Q.E.D.

Since the variables x and the constants already belong to our set of functions, the same is therefore true of functions representable analytically. The same is true of functions having different analytical expressions in different stretches and of the limits of systems of sequences of such functions.

Other functions, theoretically existent, cannot be defined by the means at our disposal and do not occur in analysis.

CHAPTER III

MONOTONE FUNCTIONS

§ 1. **Monotone functions of one variable.** A function of one variable is said to be *monotone increasing* in its interval of definition if, as t increases, it never decreases, *monotone decreasing* if it never increases.

In either case it is a *monotone function* of the variable t. A constant is obviously a monotone function, and the only monotone function of both types.

Like functions constant in stretches, monotone functions are not in general semicontinuous. A monotone ascending function is however easily seen to be upper semicontinuous on the left and lower semicontinuous on the right at every point.

A monotone function of one variable has *unique limits on either side* at each point.

It is sufficient to prove this for monotone increasing functions, and for the left-hand side.

Let t_0 be any point of the interval, and $t_1, t_2, \ldots, t_n, \ldots$ a monotone ascending sequence of numbers converging to t_0. And let $f(t)$ be the given function.

The sequence of values

$$f(t_1),\ f(t_2),\ \ldots,\ f(t_n),\ \ldots$$

is monotone ascending. We denote its limit by $f(t_0-0)$. This limit is independent of the choice of the sequence.

For suppose another sequence give a different limit; with the points of both sequences we could form a mixed sequence, still monotone ascending, which would not give a unique limit.

Now $f(t_1)$ is the lower bound of $f(t)$ in the open interval (t_1, t_0), $f(t_2)$ in (t_2, t_0), and so on. $f(t_0-0)$ is thus the limit of the lower bounds on the left, that is, the lower limit on the left. And as $f(t_0-0)$ is greater than all the values of f at points preceding t_0, it is not less than the upper limit on the left. This proves our statement.

§ 2. **Monotone functions of two variables.** Given a function of two variables $g(x, y)$ we call *simple* increment the difference

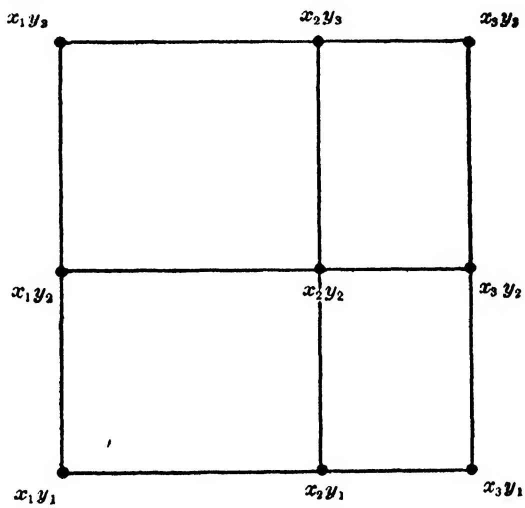
$$g(x+h, y+k) - g(x, y),$$

when neither h nor k is negative; we call *double increment* the expression*

$$\begin{bmatrix} g(x+h,\ y+k) - g(x+h,\ y) \\ -g(x,\ y+k) + g(x,\ y) \end{bmatrix}$$

when h and k are positive. The double increment is said to be taken over the interval whose endpoints are (x, y), $(x+h,\ y+k)$. It is easy to verify that the double increment possesses the following *additive property*:

If we divide a given interval into subintervals, the sum of the double increments of $g(x, y)$ over these subintervals is the double increment over the total interval.



$$\left\{\begin{aligned} & g(x_2y_2) - g(x_2y_1) - g(x_1y_2) + g(x_1y_1) \\ & + g(x_3y_2) - g(x_3y_1) - g(x_2y_2) + g(x_2y_1) \\ & + g(x_2y_3) - g(x_2y_2) - g(x_1y_3) + g(x_1y_2) \\ & + g(x_3y_3) - g(x_3y_2) - g(x_3y_2) + g(x_2y_2) \end{aligned}\right.$$
$$= g(x_3y_3) - g(x_3y_1) - g(x_1y_3) + g(x_1y_1).$$

A function of two variables is said to be *monotone increasing* in its interval of definition, if its simple *and its double*† increments are never negative. It is said to be *monotone decreasing* if these increments are never positive.

* Or, "increment with respect to y of the increment with respect to x."

† When only the simple increments are never negative, the function is sometimes called monotone increasing with respect to each variable separately. The functions defined in the text are sometimes called "entirely" monotone. If each of the three increments $g(x+h, y) - g(x, y)$, $g(x, y+k) - g(x, y)$, and the double increment have a constant sign, but this sign is not the same for the three increments, then the function is "pseudo-monotone."

For these monotone functions of two variables we have the following:

THEOREM*. A monotone function has a unique limit of approach in each open quadrant at every point.

We remark that an interval is determined when we are given any two opposite corner-points (or "diagonal" points) not necessarily endpoints. Let us denote by

$$[PQ]$$

the interval thus determined by P and Q, and by

$$\Delta[P, Q]$$

the double increment corresponding to the endpoints of $[P, Q]$. Let us suppose our function monotone increasing. It follows from our definition of such a function and from the additive property of the double increment that if Q lie in $[O, P]$, then

$$\Delta[O, Q] \leqslant \Delta[O, P].$$

When P describes a sequence of points

$$P_1, P_2, \ldots, P_n, \ldots \to O,$$

such that P_n always lies inside $[O, P_{n-1}]$, and tends to O, the double increments

$$\Delta[O, P_n]$$

form a monotone sequence and therefore have a unique limit.

This limit is independent of the sequence of points, provided they belong to the same open quadrant as the preceding sequence. For suppose another sequence give a different limit; with the points of both sequences we could form a mixed sequence with the same property as above (P_n lying inside $[O, P_{n-1}]$), which would not give a unique limit. The double increment thus has a unique limit of approach in each open quadrant. Since it consists of four terms, one of which remains constant and two of which have unique limits by the one-dimensional theorem, its fourth term also has a unique limit.

It follows immediately that this is the unique limit of approach of $g(x, y)$ in that open quadrant at O. Q.E.D.

§ 3. **Monotone functions of n variables.** Similarly to our definitions in the case of two variables, we may define simple, double, ..., n'ple increments of a function of the ensemble of n variables, corresponding to non-negative increments of these variables.

A function of the ensemble of n variables will be called *monotone increasing* if all these increments are non-negative, and *monotone decreasing* if none of them are positive.

* This is equally true (by the same argument) of a pseudo-monotone function.

We then have the theorem, proved in exactly the same manner:

A monotone function has a unique limit of approach in each open quadrant at every point*.

§ 4. **Total increment of a monotone function over a stretch.** Given a monotone function of n variables, and an open interval inside its interval of definition, the n'ple increments over closed intervals contained in it and tending to it, have a unique limit which is also their upper bound. This limit is called the *total increment* over the open interval.

A stretch which is not an interval is the limit of open intervals containing it. We call *total increment* over such a stretch the limit of the total increments corresponding to such intervals. This limit is also their lower bound.

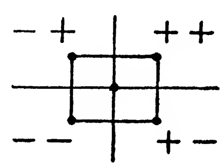
In order to make the matter clear, let us consider the two-dimensional case.

The total increment over the open interval whose endpoints are x, y and $x+h$, $y+k$ will be the expression

$$\begin{cases} g(x+h-0;\ y+k-0) - g(x+h-0;\ y+0) \\ -g(x+0;\ y+k-0) \qquad + g(x+0;\ y+0) \end{cases}$$

when $g(x+0;\ y+0)$, $g(x+0;\ y-0)$, $g(x-0;\ y+0)$, $g(x-0;\ y-0)$ denote respectively the unique limits of $g(x, y)$ in the open $(+, +)$, $(+, -)$, $(-, +)$, $(-, -)$ quadrants at x, y.

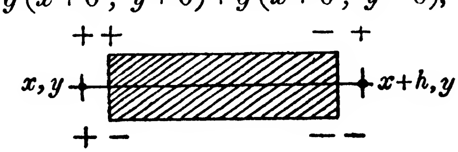
The only other kinds of stretches are points and open sides of intervals. The total increment over a point x, y will be the expression



$$g(x+0;\ y+0) - g(x-0;\ y+0)$$
$$-g(x+0;\ y-0) + g(x-0;\ y-0).$$

The total increment over the stretch x, y to $x+h$, y will be

$$g(x+h-0;\ y+0) - g(x+h-0;\ y-0)$$
$$-g(x+0;\ y+0) + g(x+0;\ y-0),$$

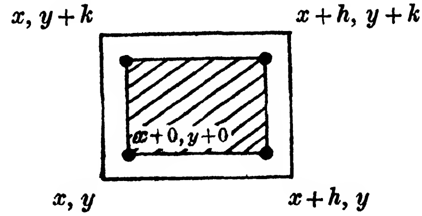


and the total increment over the stretch x, y to x, $y+k$

$$g(x+0;\ y+k-0) - g(x+0;\ y+0)$$
$$-g(x-0;\ y+k-0) + g(x-0;\ y+0).$$

* As before, it holds equally for pseudo-monotone functions.

It is convenient to think of the symbols $x-0$, $y-0$ etc. *as representing actual points* and our stretches as ordinary closed intervals with the corresponding endpoints. The total increment over a stretch is then identical in form with the double increment.



The additive property is clearly independent of the actual geometrical existence of the points of division.

It follows that *if we divide up a stretch into stretches, the total increment over the given stretch is equal to the sum of the corresponding total increments.*

§ 5. **Integration of functions constant in stretches with respect to a monotone increasing function.** Given any stretch, we define the *integral over that stretch of the constant unity* as the total increment of the given monotone increasing function over that stretch.

The integral over that stretch of a function which has there a constant value C is to be the product of this value into the total increment.

The integral over an open interval (a, b) *of a function constant in stretches* there defined is to be the sum of its integrals over the stretches of the open interval (a, b) in which it is constant.

We use the symbol

$$\int_a^b f(x)\, dg(x)$$

to denote the integral over the open interval (a, b) of the function $f(x)$ with respect to the monotone increasing function $g(x)$. The function to be integrated is called the *integrand,* the function with respect to which we integrate the *integrator.*

The integral over a stretch is defined just in the same way as the integral over an open interval, and is always expressible as such an integral possibly in a space of less dimensions.

The integrals of functions constant in stretches have the following obvious properties, which we shall call the fundamental properties of integration.

(I) If C is any constant,

$$\int_a^b C\,.\,f(x)\,.\,dg(x) = C\int_a^b f(x)\, dg(x).$$

(II) The integral of the sum of two functions is equal to the sum of their integrals

$$\int_a^b (f_1+f_2)\,dg=\int_a^b f_1\,dg+\int_a^b f_2\,dg.$$

(III) If f_1 is greater than or equal to f_2 throughout (a, b), then

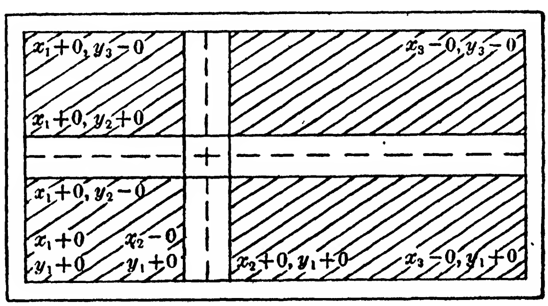
$$\int_a^b f_1\,dg \geqslant \int_a^b f_2\,dg.$$

(IV) If an interval be divided up in any manner into stretches the integral over that interval is the sum of the integrals over the partial stretches. This is really contained in (II): take $f_1=f$ in a certain stretch and zero outside, $f_2=0$ where $f_1=f$, and $f_2=f$ where $f_1=0$.

If we divide up our interval at a point we have in one dimension

$$\int_a^b fdg=\int_a^\xi fdg+\int_\xi^b fdg+f(\xi)\{g(\xi+0)-g(\xi-0)\},$$

while in two dimensions we have



$$\int_{x_1y_1}^{x_3y_3} fdg=\int_{x_1y_1}^{x_2y_2}+\int_{x_1y_2}^{x_2y_3}+\int_{x_2y_1}^{x_3y_2}+\int_{x_2y_2}^{x_3y_3}$$

$$+\int_{x_1}^{x_2} f(x, y_2)\,d\,[g(x, y_2+0)-g(x, y_2-0)]$$

$$+\int_{x_2}^{x_3} f(x, y_2)\,d\,[g(x, y_2+0)-g(x, y_2-0)]$$

$$+\int_{y_1}^{y_2} f(x_2, y)\,d\,[g(x_2+0, y)-g(x_2-0, y)]$$

$$+\int_{y_2}^{y_3} f(x_2, y)\,d\,[g(x_2+0, y)-g(x_2-0, y)]$$

$$+f(x_2, y_2)\cdot\begin{bmatrix} g(x_2+0, y_2+0)-g(x_2+0, y_2-0)\\ -g(x_2-0, y_2+0)+g(x_2-0, y_2-0)\end{bmatrix},$$

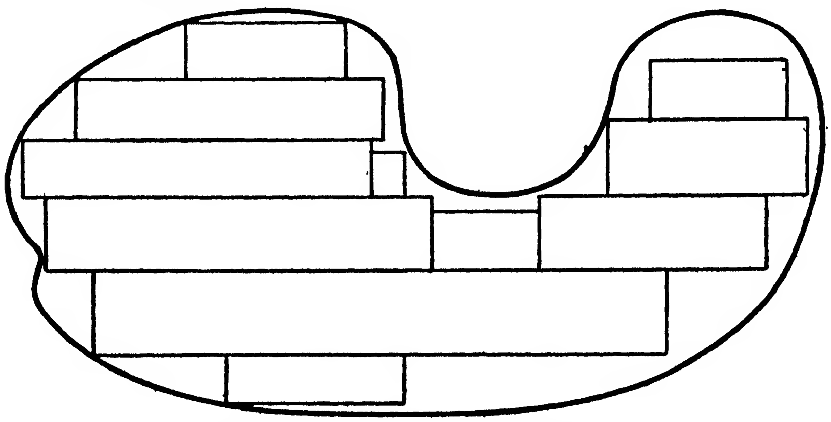
i.e. the sum of four two-dimensional integrals, four one-dimensional integrals and one, so to speak, 0-dimensional integral.

CHAPTER IV

THE INTEGRATION OF FUNCTIONS

§ 1. **Methods of the theory of integration.** The theory of integration arose from the problem of finding the area of plane curves. This led at once to the consideration of functions constant in stretches.

The original method of evaluating areas, a method known to the Greeks, consisted in attempting to construct a system of abutting rectangles such that the boundary of the area covered by them should approximate to the given curve. The problem then reduced to that of the area of rectangles.

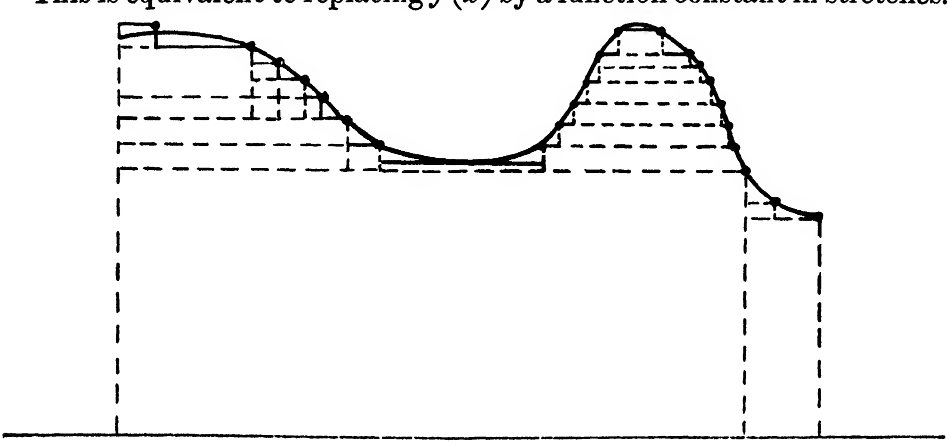


In the case of a closed curve consisting of the graph

$$y = f(x)$$

in (a, b), of the rectilinear segments $x = a$ and $x = b$, and the axis of x, it is clearly simplest to take the rectangles to have sides parallel to the axes.

This is equivalent to replacing $f(x)$ by a function constant in stretches.



The method we shall proceed to develop is a natural extension of the above. We have defined integration with respect to a monotone increasing function for functions constant in stretches; in the simplest case, in one dimension and when the integrator is the variable itself, the integral of a function constant in stretches $f(t)$,

$$\int_a^b f(t) \, . \, dt,$$

is the area of the closed curve obtained as stated above. We are therefore led to define, by successive approximations, the integral of a function which is the limit of a function constant in stretches; functions for which this definition proves valid will form the next class of integrable functions and can be used in their turn to approximate more general functions and to define their integrals; and so on. We have only to prove that at each stage the integrals of the functions by which we approximate tend to a unique limit. Such is indeed the case.

Historically the functions whose integrals with respect to a monotone increasing function were first defined were continuous. To these we easily extend integration.

THEOREM. If an ascending sequence of simple L-functions and a descending sequence of simple U-functions have the same limiting function, the limit of their integrals is the same.

We suppose all the functions defined in the closed interval (a, b). The theorem is then an immediate consequence of the theorem of bounds.

The differences between corresponding functions form a monotone descending sequence of simple U-functions

$$c_1(x), \quad c_2(x), \ldots$$

having the limit zero. By the theorem of bounds their upper bounds u_n also have the limit zero. It follows from property (III) of p. 21 that

$$\int_a^b c_n(x) \, . \, dg(x) \leqslant U_n \, . \int_a^b dg(x) \to 0.$$

The theorem then follows by property (II) of the same page.

DEFINITION. Given any continuous function we can always construct a monotone ascending sequence of simple L-functions and a monotone descending sequence of simple U-functions of which it is the limit. The limit of the integrals is, by the above theorem, independent of the choice of these sequences. (Given two pairs of sequences, we need only compare the U-sequence of one pair with the L-sequence of the other.) This limit is defined to be the integral of the given continuous function.

We have thus defined the integral of a continuous function with respect to a monotone increasing function. Such an integral is called a *Stieltjes integral.*

Stieltjes, however, only considered the case of one variable, whereas the method of monotone sequences is independent of the number of variables concerned. When the integrator is a product of these variables*, our definition of integral will be found to reduce to that of *multiple integral* such as employed for very special types of functions by *Cauchy* and *Riemann*.

§ **2. Darboux's theorem.** A modification of the method of finding areas is the following: we construct a system of abutting rectangles all internal to the given closed curve; the upper bound of the areas corresponding to all such systems is then a lower evaluation for the required area; similarly we construct a system of abutting rectangles containing the given curve; the lower bound of such areas is then an upper evaluation.

These considerations led Riemann to frame his definition of integration. Let $f(x)$ be any function defined in (a, b) and bounded in that interval. It follows from property (III) of p. 21 that the upper bound of the integrals of the functions constant in stretches which are nowhere greater than $f(x)$ is less than or equal to the lower bound of those which are nowhere less than $f(x)$. If that property is to be maintained when we extend the definition of integration to $f(x)$, the integral of $f(x)$ will have to lie between these bounds. We have thus got a lower evaluation and an upper evaluation for the integral of $f(x)$. When these are equal they give a definition of the integral of $f(x)$.

We can also get (less good) evaluations by considering a particular set of functions constant in stretches. It is not unnatural to restrict ourselves to simple functions. For the purpose of calculation the following method is convenient: we divide up the interval (a, b) in any manner into a finite number of stretches; the simple L-function, constant in these stretches, whose value in each open interval of constancy is the lower bound of $f(x)$ in the corresponding closed interval, is clearly nowhere greater than $f(x)$; let s be its integral. The upper bound of all such numbers s we shall call *Darboux's lower evaluation* of the integral of $f(x)$ with respect to $g(x)$. Similarly the simple U-function obtained by doing the same with the upper bounds of $f(x)$ is nowhere less than $f(x)$; let S be its integral. The lower bound of all such numbers S we shall call *Darboux's upper evaluation* of the integral.

When these two bounds are equal they give *Riemann's definition of the integral of* $f(x)$; this definition is however quite insufficient for

* The total increment of the integrator over any stretch is then the area, or volume or etc. of that stretch.

our purposes; in fact, unless the integrator is continuous, Darboux's evaluations do not even coincide in the case of simple functions.

Given a subdivision of the interval (a, b) we shall say of another subdivision that it is consecutive to it, if it is got by dividing up the stretches of the first subdivision.

The simple L-function corresponding to the consecutive subdivision is obviously nowhere less than the former one. Corresponding to a system of consecutive subdivisions we thus have a monotone ascending sequence of simple L-functions. Darboux was able to prove that, in the case of a continuous integrator, the *integrals of this sequence tend to his lower evaluation*, provided the length of the longest stretch of the subdivision tends to zero. This theorem is equally true when the integrator is not assumed to be continuous*.

For our purpose it is sufficient to prove Darboux's theorem in the case where $f(x)$ is an L-function.

Given any function constant in stretches, a simple L-function and a simple U-function are uniquely determined coinciding with it in those stretches of constancy which are open intervals. We shall call them the *related* simple L- and U-functions.

If $f(x)$ is an L-function bounded below (as we are only concerned with the lower evaluation, f need not be bounded above) the simple L-functions of Darboux's theorem tend to f, and so do also their related simple U-functions (p. 10). We have only to prove the following lemma.

LEMMA. If two monotone ascending sequences of simple L-functions and also the related simple U-functions all have the same limiting function, then the integrals of the simple L-functions of the two sequences have the same limit.

Let $a_n(x)$, $b_n(x)$ be the generic terms of the two sequences of simple L-functions, $a_n'(x)$, $b_n'(x)$ the related simple U-functions.

Let $f(x)$ be the limiting function.

Suppose all these functions defined in the closed interval (a, b).

Let $b_N(x)$ be any one of the $b_n(x)$; let d be any one of the open intervals in which it is constant and D the interval got by closing d.

* A similar proposition holds for the upper evaluation.

As a matter of fact both follow from the theorems of this chapter. For the limit of the given ascending sequence of simple L-functions is at any point the minimum limit at the point. Hence it is sufficient to prove, as we do further on, that if two monotone ascending sequences of simple L-functions have the same limiting function, then the limits of their integrals coincide.

In D, $f(x)$ is not less than b_N' and its lower bound not less than the lower bound of b_N', which coincides with the upper bound of b_N, being the common constant value in d.

By the theorem of bounds applied to the first sequence, if m denote any number smaller than the lower bound of $f(x)$ in D, we can find an n_0 from and after which, in D, the lower bound of $a_n(x)$ is greater than or equal to m. Since we may choose for m the upper bound of b_N in D, minus e, we have, a fortiori, throughout D,

$$a_n(x) \geqslant b_N(x) - e$$

for all n greater than n_0. Corresponding to each D, we have a similar inequality.

If n_1 denote the greatest of the corresponding n_0's, we have, throughout (a, b),

$$a_n(x) \geqslant b_N(x) - e$$

for all n greater than n_1.

By properties (I) and (II), remembering that e is arbitrary, we deduce, passing to the limit

$$\lim_{n=\infty} \int_a^b a_n(x)\, dg(x) \geqslant \lim_{n=\infty} \int_a^b b_n(x)\, dg(x).$$

But, reversing the rôles of the sequences, we get similarly,

$$\lim_{n=\infty} \int_a^b a_n(x)\, dg(x) \leqslant \lim_{n=\infty} \int_a^b b_n(x)\, dg(x).$$

Hence

$$\lim_{n=\infty} \int_a^b a_n(x)\, dg(x) = \lim_{n=\infty} \int_a^b b_n(x)\, dg(x). \qquad \text{Q.E.D.}$$

Corollary I. The above lemma also holds for integrals over a stretch.

(For an integral over a stretch which is not an open interval is expressible as an integral over an open interval of a space of less dimensions.)

Corollary II. If the limiting function is a simple L-function the limit of the integrals is the integral of the limiting function.

Case 1. f is a constant. We take for one of our sequences the repetition of the function itself.

Case 2. f is any simple function. We divide up the interval into the stretches where f is constant. In each of these, by case 1, the limit of the integrals is equal to the integral of the limiting function. By property (IV), p. 21, the proposition follows by addition.

Theorem. (Darboux's theorem for an L-function.) Darboux's lower evaluation of the integral of an L-function bounded below is the unique

limit of the numbers s corresponding to any system of consecutive subdivisions of (a, b) when the length of the longest stretch of the subdivision tends to zero.

Let f_n be the simple L-function corresponding to the nth subdivision of (a, b). By the method employed to prove the theorem of p. 10, f_n, and also the related simple U-function f_n', tend to the given L-function f. Now consider any two systems of consecutive subdivisions. The corresponding two sequences of simple L-functions fulfil the conditions of our lemma, and therefore the two monotone ascending sequences of numbers s have the same limit, which is their common upper bound. This limit is therefore unique and is the upper bound of all the numbers s.

§ 3. **The integration of semicontinuous functions.** The preceding theorem enables us to define integration for any L-function bounded below, in accordance with the methods of § 1. Since Darboux's lower evaluation is the limit of the integrals of certain monotone sequences of simple L-functions which tend to $f(x)$, it gives a natural definition of the integral of $f(x)$.

Definition. The integral of an L-function bounded below is Darboux's lower evaluation.

We shall say a sequence of functions is integrable term by term if the limit of the integrals (with respect to a given monotone increasing function) equals the integral of the limiting function.

Theorem. Every monotone ascending sequence of L-functions, bounded below, is integrable term by term.

Let $f_1, f_2, \ldots, f_n, \ldots$ be the sequence and f the limiting function. The integral of f_1 is defined by means of the auxiliary monotone ascending sequence of simple L-functions

$$f_{11}, f_{12}, f_{13}, \cdots, f_{1n}, \cdots$$

which tend to f_1, as do also the related simple U-functions.

Similarly f_2 is the limit of an auxiliary ascending sequence of simple L-functions

$$f_{21}, f_{22}, f_{23}, \cdots, f_{2n}, \cdots$$

and of the related simple U-functions. Since f_1 is nowhere greater than f_2 it is clear that the latter is also the limit of the simple L-function which is at every point equal to the greater of f_{1n} and f_{2n}, and also of

the simple U-function which is the greater of their related simple U-functions. We can consequently always so arrange that

$$f_{2n} \geqslant f_{1n} \text{ for all } n.$$

Proceeding in this way, we get, as on p. 26, a monotone ascending sequence of simple L-functions

$$f_{11}, f_{22}, f_{33}, \ldots$$

which, and also the related simple U-functions, tend to f; they therefore define the integral of f.

Now this sequence has all its terms after the mth greater than the corresponding terms of the auxiliary sequence of f_m. Hence

$$\int_a^b f\,dg \geqslant \int_a^b f_m dg \text{ for all } m.$$

But we also have, for all m,

$$\int_a^b f_m dg \geqslant \int_a^b f_{mm} dg.$$

The theorem follows by making $m \to \infty$.

In particular the integral of an L-function bounded below is the limit of the integrals of any monotone ascending sequence of simple L-functions of which it is the limiting function.

Remembering that change of sign transfers a type to the twin type, we see that similarly we may define the *integral of any U-function bounded above** as the limit of the integrals of any monotone descending sequence of simple U-functions which has it as limiting function. Also we have the theorem that every monotone descending sequence of U-functions is integrable term by term.

The properties of the integrals of functions constant in stretches are at once extensible to those of semicontinuous functions.

I. If C is any constant, then $\int Cfdg = C\int fdg$.

For the generating simple functions of $C.f$ are C times those of f.

II. The integral of the sum of two functions is equal to the sum of their integrals.

This presupposes that the sum is semicontinuous and that the sum of the integrals has a meaning.

* In the case of functions of both types, continuous functions, we have already seen that the two definitions coincide; it is also easily seen independently, as the difference of the two sequences is a monotone descending sequence of simple U-functions tending to zero, which as we know is integrable term by term.

If both functions are L-functions, their sum is the limit of the sum of their auxiliary sequences. The other cases reduce to this, because $-L=U$.

III. If f_1 is greater than or equal to f_2 throughout (a, b), then

$$\int_a^b f_1 dg \geqslant \int_a^b f_2 dg.$$

If both are L-functions we can so arrange that the generating functions of the latter are all less than or equal to the corresponding ones of the former.

If one is an L-function and the other a U-function, the difference of their generating sequences tends to a nowhere negative limiting function, which can also be generated by nowhere negative simple functions.

IV. If the interval be divided up in any manner into stretches the integral over that interval is the sum of the integrals over the partial stretches.

It follows immediately from (III) that the integrals of the simple U-functions of paragraph 2 (which lie between the corresponding simple L-functions and the limiting L-function) tend to that of the limiting function. From this we deduce that, *given any L-function bounded below, its integral is equal to the upper bound of the integrals of the U-functions bounded above nowhere greater than it.*

CHAPTER V

THE INTEGRATION OF FUNCTIONS (*continued*)

§ 1. **Final extension of integration.** Besides not being applicable to simple functions, Riemann's definition has the great disadvantage that it will in general fail in the case of the limiting function of a sequence whose terms are integrable according to his definition.

It is a remarkable fact that both these drawbacks disappear if we apply the very same method at the next stage.

DEFINITION. Given any function $f(x)$ whatever, bounded or unbounded, the upper bound of the integrals of the U-functions bounded above which nowhere exceed $f(x)$ is called the *lower integral* of $f(x)$; the lower bound of the integrals of all the L-functions bounded below which are nowhere less than $f(x)$ is called the *upper integral* of $f(x)$*.

* N.B. If we permute the letters U and L we get Darboux's lower and upper evaluations.

DEFINITION. If its upper and lower integrals are equal, $f(x)$ is said to *possess an integral* (finite or infinite) which is their common value. If the integral is finite, $f(x)$ is said to be *integrable.*

FUNDAMENTAL THEOREM. The limiting function of a monotone sequence of integrable functions possesses an integral (finite or infinite).

Let $f_1, f_2, \ldots, f_n, \ldots$ be the given sequence and let f be its limiting function.

It follows from our definition of integral that there exist, corresponding to each function f_n of the sequence, U-functions bounded above nowhere greater than f_n and L-functions bounded below nowhere less than f_n whose integrals differ from that of f_n by less than $e \,.\, 2^{-n-1}$. Let $A_n(x)$ and $B_n(x)$ be any two of these, such that

$$A_n \geqslant f_n \geqslant B_n.$$

Let us suppose the given sequence was monotone ascending. We assert that any pair of the sequences A_n and B_n can be replaced by a pair of monotone ascending sequences a_n and b_n, such that

$$a_n \geqslant f_n \geqslant b_n,$$

$$\int a_n(x)\,dg(x) - \int b_n(x)\,dg(x) \leqslant e.$$

We need only choose $a_1 = A_1$, $b_1 = B_1$ and then define by recurrence a_n to be the function which is equal at every point to the greater of a_{n-1} and A_n, and similarly b_n the greater of b_{n-1} and B_n. By the fundamental property (ii) of p. 9, a_n is an L-function bounded below and b_n a U-function bounded above.

Further, we have at every point

$$\begin{aligned} a_2(x) - b_2(x) \leqslant \text{ either } & A_1(x) - B_1(x) \\ \text{or } \; & A_2(x) - B_2(x). \\ \therefore \leqslant (A_1(x) - B_1(x)) &+ (A_2(x) - B_2(x)). \end{aligned}$$

Hence

$$\int a_2(x)\,dg(x) - \int b_2(x)\,dg(x) \leqslant \left\{ \begin{aligned} &\int A_1(x)\,dg(x) - \int B_1(x)\,dg(x) \\ &+ \int A_2(x)\,dg(x) - \int B_2(x)\,dg(x), \end{aligned} \right.$$

$$\therefore \leqslant e \,.\, (\tfrac{1}{2} + \tfrac{1}{4}).$$

Similarly

$$a_n(x) - b_n(x) \leqslant \{a_{n-1}(x) - b_{n-1}(x)\} + \{A_n(x) - B_n(x)\}$$

and $$\int a_n(x)\,dg(x) - \int b_n(x)\,dg(x) \leqslant e\,.\left(\frac{1}{2}+\frac{1}{4}+\dots+\frac{1}{2^n}\right)$$
$$\leqslant e.$$

Now the limiting function of the sequence a_n is an L-function bounded below nowhere less than f and whose integral we have seen to be equal to the limit of those of its generating functions. Also $\int b_n dg$ remains always less than or equal to the upper bound of the integrals of the U-functions bounded below nowhere greater than f.

It follows immediately since e is arbitrary that the theorem is true.

THEOREM. A monotone sequence of integrable functions may be integrated term by term.

This follows immediately from the preceding argument since both the limit of the integrals and also the integral of the limiting function lie between the limit of $\int a_n dg$ and that of $\int b_n dg$, and since these limits differ by less than e.

THEOREM. A function obtainable by monotone sequences is either integrable or else both its upper and its lower integrals are infinite.

Suppose the lower integral is finite, there is then a U-function bounded above, having a finite integral, and nowhere greater than the given function f. It follows (from the second of the three fundamental properties of class) that the latter is obtainable by monotone sequences of functions greater than the U-function.

Now let $f_n(x)$ be the function which is at every point the lesser of the given function and of the constant n, this function is of the same type as f. Also it lies between the constant n and a certain integrable U-function; and we may therefore suppose this property possessed by all the functions of the system of monotone sequences defining f_n. It follows immediately by repeated application of the fundamental theorem, remembering that by the theorem just proved the integrals obtained at each stage are finite, that $f_n(x)$ is integrable.

Our given function therefore possesses an integral, by that same fundamental theorem. This integral cannot be infinite by hypothesis. The theorem follows.

For integrable functions we easily establish the four properties of

pp. 20, 21. The proof is indeed but a repetition of the argument used in the case of semicontinuous functions.

(I) $\int Cf\,dg = C.\int f\,dg$,

(II) $\int (f_1 + f_2)\,dg = \int f_1\,dg + \int f_2\,dg$,

(III) If $f_1 \geqslant f_2$ then $\int_a^b f_1\,dg \geqslant \int_a^b f_2\,dg$,

(IV) The $\int$ over an open interval = sum of $\int$'s over parts.

Let f be any integrable function (its integrability depends on the integrator; all functions are integrable with respect to a constant and have their integrals equal to zero).

Then f lies between two integrable semicontinuous functions, i.e. a U-function bounded above and an L-function bounded below. By adding suitable constants to them, we see that we may take the U-function to be negative and the L-function to be positive at every point, without affecting their integrability.

On the other hand it follows immediately from the last theorem of the preceding paragraph that *any function which lies between two integrable functions is itself integrable*, if it is one which can be obtained by monotone sequences.

This is in particular the case of the function equal to f, where f is positive, and zero elsewhere; and also of the function equal to f, where f is negative, and zero elsewhere. Changing the sign of the latter function and adding it to the other we get $|f|$. Consequently $|f|$ is integrable. *Thus if f is integrable so is $|f|$.*

For this reason a function which is integrable with respect to $g(x)$ according to our definition is said to possess an *absolutely convergent integral* with respect to $g(x)$.

Non-integrable functions fall into two classes: those possessing an integral (which is then infinite) and those whose upper integral is $+\infty$ and whose lower integral is $-\infty$.

For this second class of functions our definition fails entirely. It is however sometimes convenient to define for them a generalised integral with respect to $g(x)$. Under what conditions this is possible we do not propose to investigate here. That it is not always possible is easy to see; we need only take the case of a function which is $+\infty$ throughout one interval and $-\infty$ throughout another.

N.B. If D is a stretch

(1) over which the total increment of g vanishes, or

(2) throughout which f is zero,

then
$$\int_D f dg = 0$$
(for this is obvious by definition if f is constant in stretches; it is therefore true for semicontinuous functions and so on by induction for all functions mathematically definable).

From (2) and the fundamental property of integration II, it follows at once that assigning arbitrary values to a function on the boundary of the interval (a, b) will not affect its integral over that interval.

§2. Indefinite, parametric and repeated integration.

A function which is integrable in (a, b), is also integrable in every open interval contained in (a, b).

For let D be such an interval and f_1 the function equal to f in D and zero outside. Then
$$0 \leqslant |f_1| \leqslant |f|$$
at every point. Hence f_1 lies between the two integrable functions and, since it is obtainable by monotone sequences, it is integrable in (a, b).

Further,
$$\int_a^b f_1 dg = \int_D f dg.$$

In particular, if ξ is any point of (a, b),
$$\int_a^\xi f dg = F(\xi), \text{ say,}$$
exists and is finite.

The function $F(x)$ is called the *indefinite integral* of f with respect to g. It is obvious that *if $f \geqslant 0$, F is monotone increasing.* For a simple increment of F is a sum of integrals over stretches and cannot be negative; a double increment of F is an increment of an increment, i.e. a simple increment of a sum of integrals, i.e. a sum of integrals, and therefore not negative. Similarly, all the increments of F simple, double, ..., $n$ple, are sums of integrals of f and consequently non-negative.

Since any integrable function can be regarded as the difference of two positive integrable functions, *an indefinite integral is the difference of two monotone increasing functions.*

We assert that, moreover, F is continuous at the endpoints of (a, b), i.e., that the unique limit of F in the (all +)-quadrant at a is zero, the unique limit in the (all −)-quadrant at b is $F(b)$.

Let a_n, b_n be any two points of the closed interval (a, b) such that (a_n, b_n) forms an interval, and let them describe two sequences of points such that

$$a_n \text{ lie inside } (a, a_{n-1}) \text{ and } a_n \to a$$
$$b_n \text{ ,, } \quad \text{ ,, } \quad (b_{n-1}, b) \text{ and } b_n \to b.$$

Let us suppose the integrand f is positive and let f be the function equal to f in (a, b) and zero elsewhere. By the remark at the end of § 1, we may clearly suppose f equal to zero outside the open interval (a, b), if it is even defined.

Then, as n increases, f_n forms a monotone sequence of integrable functions whose limit is f. Therefore

$$\int_a^b f_n dg \to \int_a^b f dg,$$

i.e.
$$\int_{a_n}^{b_n} f dg \to \int_a^b f dg.$$

This proves our statement in this case.

If f is the difference of two positive integrable functions, the result now follows at once.

It immediately follows that the indefinite integral F of f with respect to g has at most the discontinuities of $g(x)$. If $f=0$ where g is discontinuous, F is continuous everywhere.

An indefinite integral is thus in our scheme of functions a very simple function indeed.

Let us now suppose that instead of one function f and one function g, we are considering at every point x *a set of pairs of functions* $f(x)$ and $g(x)$, the f's being all obtainable by monotone sequences and the g's being further all monotone increasing functions. This set might be, for example, the succession

$$f_1, g_1; \quad f_2, g_2; \quad \ldots; \quad f_n, g_n; \quad \ldots$$

We might then suppose that, besides the coordinates which represent the point x, we have a further coordinate p representing the index n, for example, its inverse $1/n$.

In this new space we should have the two functions

$$f(x, p), \quad g(x, p)$$

defined at all points (x, p) whose p-coordinate corresponds to an integral value of n, and we might, for example, agree to give to both the value zero elsewhere.

More generally, we may suppose that we are dealing with two functions

$$f(x, p), \quad g(x, p),$$

of *two sets of variables* x and p, and that when p is fixed, g is monotone increasing with respect to x. If we now keep p fixed, and form the indefinite integral

$$\int^x f dg,$$

we shall have a new function of the two sets of variables x and p,

$$F(x, p) = \int^x_{p \text{ fixed}} f(x, p)\, dg(x, p).$$

This process is called *parametric integration.*

We may remark that $F(x, p)$ is obtainable by monotone sequences, for integration is a limiting process and we have seen that such a process performed on functions obtainable by monotone sequences, leads to the same kind of function. In fact, if $f(x, p)$ is obtainable by monotone sequences, starting with functions constant in stretches, F is so too starting with functions constant in stretches with respect to p and monotone increasing with respect to x, and these are themselves obtainable by monotone sequences starting with functions constant in stretches with respect to (x, p).

In particular, we may apply to F the process of integration with respect to a monotone increasing function of p (keeping x fixed). We then obtain what may be called a *repeated indefinite integral.*

While integrating with respect to x, we may also keep fixed the interval over which we integrate; for example, it might be the interval (a, b); we should then obtain by integration the same result as by substituting b for x in $F(x, p)$, i.e. $F(b, p)$.

More generally, we may suppose the interval a to x over which we integrate keeping p fixed, to depend on p. Integration would then lead to the same result as substituting for the variables x in F certain functions of p, i.e. we should get

$$F(x(p), p).$$

Integration with respect to a monotone increasing function of the variables p will then give us the most general kinds of *repeated definite integral.*

§ 3. Term-by-term integration of sequences and successions.

THEOREM I. If a monotone sequence of integrable functions whose limit is integrable be integrated term by term, then the integrated sequence converges uniformly to the indefinite integral of the limiting function.

(In the case when the integrator g is continuous, this is a corollary of Appendix, 1, p. 46.)

CASE 1. The functions are all positive. Then

$$\left|\int_a^x f_n\,dg - \int_a^x f\,dg\right| \leqslant \left|\int_a^b f_n\,dg - \int_a^b f\,dg\right|.$$

We can find an n from and after which the righthand side is less than e; the same n makes the lefthand side less than e.

CASE 2. General case.

$$f_n = a_n - b_n; \quad f = a - b,$$

where a_n, b_n, a, b, are positive functions and we can so arrange that a_n tends to a, b_n to b.

By case 1, we can find an n from and after which both

$$\left|\int_a^x a_n\,dg - \int_a^x a\,dg\right| \quad \text{and} \quad \left|\int_a^x b_n\,dg - \int_a^x b\,dg\right|$$

are less than $e/2$. Then, from and after that n,

$$\begin{aligned}\left|\int_a^x f_n\,dg - \int_a^x f\,dg\right| &= \left|\left(\int_a^x a_n\,dg - \int_a^x a\,dg\right) - \left(\int_a^x b_n\,dg - \int_a^x b\,dg\right)\right| \\ &\leqslant \left|\int_a^x a_n\,dg - \int_a^x a\,dg\right| + \left|\int_a^x b_n\,dg - \int_a^x b\,dg\right| \\ &< e.\end{aligned}$$

This proves the theorem.

THEOREM II. If a sequence of functions whose bounding functions are integrable be integrated term by term, the integrated sequence converges uniformly to the integral of the limiting function.

(This contains the preceding as a particular case.)

f is the limit of a monotone ascending sequence a_n, and of a monotone descending sequence b_n, of integrable functions, such that

$$a_n \leqslant f_n \leqslant b_n$$

at every point. Determine N so that both

$$\left(\int_a^x a_n\,dg - \int_a^x f\,dg\right) \quad \text{and} \quad \left(\int_a^x b_n\,dg - \int_a^x f\,dg\right)$$

are less than e in absolute value.

$$\int_a^x f_n\,dg - \int_a^x f\,dg$$

lies between them, i.e., a fortiori, is less than e in absolute value. Q.E.D.

THEOREM III. If a succession of functions whose bounding functions are integrable, and whose upper and lower functions have the same integral be integrated term by term, the integrated succession will cease to oscillate, and will converge uniformly to the common value of the integrals of the upper and lower functions.

(This again contains both the preceding theorems as particular cases.)

We have
$$\int_a^x a_n\, dg \leqslant \int_a^x f_n\, dg \leqslant \int_a^x b_n\, dg$$
and both the extreme left and the extreme righthand side converge uniformly to their common limit. Q.E.D.

THEOREM IV. Given a succession of functions f_n such that
$$\int f_n dg \rightarrow \int f\, dg$$
uniformly in every interval, if, further, f_n is in absolute value less than a fixed constant, and ϕ is any integrable function, then
$$\int \phi f_n\, dg \rightarrow \int \phi f\, dg \text{ uniformly.}$$

We first show that if ϕ is any integrable function, there are functions constant in stretches ϕ_m such that
$$\int_a^b \left| \phi - \phi_m \right| dg < \frac{1}{m}.$$

From the integrability of ϕ, we deduce the existence of an L-function ϕ_m' bounded below and nowhere less than ϕ, such that
$$\int_a^b (\phi_m' - \phi)\, dg < \frac{1}{2m}.$$

This L-function is the limit of an ascending sequence of simple L-functions. There therefore exists a simple L-function ϕ_m nowhere greater than ϕ_m' and such that
$$\int_a^b (\phi_m' - \phi_m)\, dg < \frac{1}{2m}.$$

But
$$\begin{aligned} |\phi - \phi_m| &\leqslant |\phi - \phi_m'| + |\phi_m' - \phi_m| \\ &= (\phi_m' - \phi) + (\phi_m' - \phi_m). \end{aligned}$$

Hence
$$\int_a^b \left| \phi - \phi_m \right| dg < \frac{1}{2m} + \frac{1}{2m} = \frac{1}{m}.$$

The theorem to be proved now follows at once. The absolute values of f_n are less than a fixed constant M, say. Now
$$\int \phi f_n\, dg = \int \phi_m f_n\, dg - \int f_n (\phi_m - \phi)\, dg;$$

therefore

$$\int \phi f_n \, dg - \int \phi f \, dg = \int \phi_m (f_n - f) \, dg$$
$$+ \int f(\phi_m - \phi) - \int f_n (\phi_m - \phi) \, dg.$$

Since ϕ_m is constant in stretches, the first integral on the right tends uniformly to zero if m is fixed. The absolute value of each of the two last integrals is less than M/m, a quantity which can be made as small as we please by making m large enough. The theorem follows at once.

N.B. The fact that the succession was bounded was essential to the proof. In the next theorem, this is replaced by a more general condition; but ϕ is bounded.

THEOREM V. If

$$\int f_n \, dg \rightarrow \int f \, dg$$

uniformly in every interval; if, further, $|f_n|$ remains less than an integrable function ψ, and if ϕ is any bounded function, then

$$\int \phi f_n \, dg \rightarrow \int \phi f \, dg \text{ uniformly.}$$

Let $\psi_{(M)}$ be the function which is equal to ψ where ψ is greater than the constant M, and which is equal to M elsewhere. We first prove that we can always so determine M that

$$\int (\psi_{(M)} - M) \, dg < e.$$

The functions $\{\psi + M - \psi_{(M)}\}$ form, as M increases to infinity, a monotone ascending sequence of functions whose limit is ψ, for each of them is equal to ψ where ψ is $< M$, and equal to M elsewhere. We can therefore determine M so that

$$\int \{\psi + M - \psi_{(M)}\} \, dg > \int \psi \, dg - e$$

by the theorem of term by term integration of a monotone sequence. M so determined will do.

By the fundamental property of integration III, we then have, a fortiori, for all n,

$$\int (|f_n|_{(M)} - M) \, dg < e.$$

It is clearly no restriction to suppose that, also

$$(|f|_{(M)} - M) \, dg < e,$$

which certainly holds if $|f|$ is smaller than ψ. We now proceed as in the last theorem. Let B be the maximum of ϕ, and let ϕ_m be a function constant in stretches nowhere greater than B in absolute value, such that

$$\int \left| \phi - \phi_m \right| dg < \frac{e}{M}.$$

Then, as before,

$$\int \phi f_n \, dg - \int \phi f \, dg = \int f (\phi_m - \phi) \, dg - \int f_n (\phi_m - \phi) \, dg + \int \phi_m (f_n - f) \, dg \qquad \ldots\ldots(1).$$

Now
$$\left| \int f_n (\phi_m - \phi) \, dg \right| \leqq \int |f_n| \left| \phi_m - \phi \right| dg$$
$$< \int |f_n|_{(M)} \cdot \left| \phi_m - \phi \right| dg$$
$$\leqq M \int \left| \phi_m - \phi \right| dg + 2B \int (|f_n|_{(M)} - M) \, dg$$
$$< e \,.\, (1 + 2B).$$

Similarly,
$$\left| \int f (\phi_m - \phi) \, dg \right| < e \,.\, (1 + 2B).$$

Further, since ϕ_m is constant in stretches, we have uniformly,

$$\underset{n=\infty}{Lt} \int \phi_m (f_n - f) \, dg = 0.$$

Hence, since e is arbitrary, the lefthand side of (1) differs as little as we please from zero. Q.E.D.

§ 4. Theorems concerning the integrator.

It is convenient to define integration with respect to a function *which is the sum of two monotone functions* by means of the formula

$$\int f \, d(g_1 + g_2) = \int f \, dg_1 + \int f \, dg_2.$$

This formula holds when $g_1 + g_2$ are monotone in the same sense: this is obvious when f is a constant and is therefore equally true when f is constant in stretches. Also if it is true of the generic term of a monotone sequence, it is true of the limiting function. (It is assumed that the righthand side has a meaning.)

It follows immediately from the validity of the formula in this case that when the functions g_1, g_2 are monotone in opposite senses, the righthand side remains unaltered when g_1 and g_2 are replaced by g_1', g_2'

where $g_1+g_2=g_1'+g_2'$. In fact we have, assuming g_1, g_1' monotone in the same sense,

$$\left\{\int f\,dg_1+\int f\,dg_2\right\}-\left\{\int f\,dg_1'+\int f\,dg_2'\right\}=\int f\,d(g_1-g_2')-\int f\,d(g_1'-g_2)=0.$$

The definition is therefore consistent, that is to say, independent of the choice of g_1 and g_2, dependent only on their sum (g_1+g_2).

Theorems involving variation of the integrator are difficult unless they can be deduced from theorems involving only variation of the integrand. In this way we are able to prove the following theorems.

THEOREM I. If u, v and uv are integrable with respect to g, and

$$U=\int_a^x u\,dg, \quad V=\int_a^x v\,dg,$$

then
$$\int_a^x uv\,dg=\int_a^x u\,dV=\int_a^x v\,dU.$$

It is no restriction to suppose v *positive* integrable as it is the difference of two such functions.

Now if u is constant, the theorem is obvious. If u is constant in stretches, the theorem, true in each of these, is by the fundamental property IV true in their sum. Finally, if u is the limit of a monotone sequence u_n, uv is the limit of the *monotone* sequence $u_n \,.\, v$ (monotone, since v is positive). By the theorem on term by term integration of a monotone sequence, if the theorem is true for u_n, it is equally true for u. This proves the theorem by induction for all functions mathematically definable.

THEOREM II. If the integrator g is a function of the two sets of variables x and y, while f depends on the variables x only, then

$$\int_{00}^{ab} f\,dg=\int_0^a f\,d\left\{\int_0^b \underset{x \text{ is fixed}}{dg}\right\}.$$

This is proved in exactly the same way as the preceding theorem. It is obvious for a constant, hence true for a simple function, and by the theorem on term by term integration of a monotone sequence follows in general.

THEOREM III. If for a constant t, $g(x,t)$ is the difference of two monotone increasing functions of x, and if $h(t)$ is the difference of two monotone increasing functions of t, then

$$\int^t \underset{x \text{ fixed}}{dh(t)}\left\{\int^x f(x)\,\underset{t \text{ fixed}}{dg(x,t)}\right\}=\int^x f(x)\,d\left\{\int^t \underset{x \text{ fixed}}{g(x,t)}\,dh(t)\right\}.$$

It is again sufficient to prove this for the case when $f(x)$ is constant, that is to say it suffices to prove that

$$\int^{t} dh(t)_{x \text{ fixed}} \left\{ \int^{x} dg\underset{t \text{ fixed}}{(x, t)} \right\} = \int^{x} d_{t \text{ fixed}} \left\{ \int^{t} g\underset{x \text{ fixed}}{(x, t)}\, dh(t) \right\}.$$

Now, if x is one single variable,

$$\int_{a}^{x} dg\underset{t \text{ fixed}}{(x, t)} = g(x, t) - g(a, t).$$

If we integrate with respect to $h(t)$, we get, on the right,

$$\int g\underset{x \text{ fixed}}{(x, t)}\, dh(t) - \int g(a, t)\, dh(t).$$

If x stands for two variables, say ξ and η, then

$$\int_{a}^{x} dg\underset{t \text{ fixed}}{(x, t)} = \begin{bmatrix} g\{(\xi - 0, \eta - 0), t\} + g\{(\alpha + 0, \beta + 0), t\} \\ -g\{(\xi - 0, \beta + 0), t\} - g\{(\alpha + 0, \eta - 0)\ t\} \end{bmatrix}$$

and if we integrate with respect to $h(t)$ we again get the same as

$$\int^{x} d_{t \text{ fixed}} \left\{ \int^{t} g\underset{x \text{ fixed}}{(x, t)}\, dh(t) \right\}.$$

Similarly, the formula holds if x stands for any number of variables. This proves the theorem when $f(x)$ is a constant. It follows at once that it is true for f constant in stretches. Now let

$$f_1, f_2, \ldots, f_n, \ldots \rightarrow f$$

be a monotone sequence of functions, converging to f. Then

$$\int f\, dG = \operatorname{Lim} \int f_n\, dG$$

for every G, difference of two monotone increasing functions. In particular

$$q(t) = \int f(x)\, dg\underset{t \text{ fixed}}{(x, t)} = \operatorname{Lim} q_n(t) = \int f_n(x)\, dg\underset{t \text{ fixed}}{(x, t)}.$$

Also, $q_n(t)$ forms as n increases a monotone sequence (g being taken to be monotone increasing, for brevity). Therefore,

$$\int q(t)\, dh(t) = \operatorname{Lim} \int q_n(t)\, dh(t).$$

But since f_n may be integrated term by term with respect to

$$\left[\int^{t} g\underset{x \text{ fixed}}{(x, t)}\, dh(t) \right],$$

it follows that if the theorem holds for f_n, it does also for f.

Theorem IV.

$$\int^{y} \left\{ \int^{x} f\underset{y \text{ fixed}}{(x, y)}\, dg(x) \right\} dh(y) = \int^{xy} f(x, y)\, d\left\{ g(x) \cdot h(y) \right\}.$$

The proof is so similar to the preceding ones that it may be left as an exercise to the student.

APPENDIX

1. *Continuity and uniform convergence.*

We saw that a function is said to be continuous at a point where it is both upper and lower semicontinuous and assumes a finite value. Hence, if $f(x)$ is continuous at a point x_0, there is an interval surrounding x_0 throughout which

$$|f(x)-f(x_0)| < e.$$

This is the e-definition of continuity at a point, and is due to *Cauchy.* Furthermore, if we are given any sequence of points contained in the interval of definition, and converging to x_0 as limiting point, the corresponding sequence of values of f will have $f(x_0)$ as unique limit. This is *Heine's* definition of continuity. The three definitions of continuity are absolutely equivalent, provided in the case of the third we add the condition that $f(x_0)$ is finite. It is clear that both Cauchy's and Heine's definition can be deduced from ours. $f(x_0)$ is the unique limit of the upper and of the lower bounds of f in the neighbourhood of x_0. We can therefore determine an interval containing x_0 in which the difference between the upper and lower bounds of f is less than e. Throughout that interval, we shall have, a fortiori,

$$|f(x)-f(x_0)| < e.$$

On the other hand, if we are given a sequence of points

$$x_1 \dots x_n \dots$$

converging to x_0 as limiting point, then the corresponding sequence of values of f lies between the two sequences of the upper and lower bounds of f in the intervals

$$(x_1 x_0) \dots (x_n x_0) \dots$$

which shrink up to x_0.

Conversely our definition may be deduced from any of the two others. If we have determined an interval surrounding x_0 throughout which

$$|f(x)-f(x_0)| < \tfrac{1}{2}e,$$

then the same inequality holds if we replace $f(x)$ by its upper or its lower bound; adding the two inequalities thus obtained, it follows that these bounds differ by less than e.

On the other hand, if Heine's definition is fulfilled, consider any sequence of intervals shrinking up to x,

$$w_1, w_2, w_3, \dots.$$

Let x_1 be any point of w_1, where f assumes a value differing from its upper bound by less than e, x_n any point of w_n where f_n assumes a value differing by less than e/n from its upper bound in w_n; then the points*

$$x_1, x_2, x_3, \dots$$

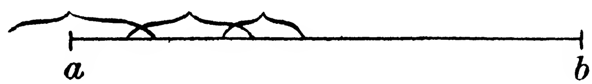
* This involves the so-called Zermelo axiom. A similar argument was used to prove the fundamental theorem of integration on p. 30 but might have been avoided.

form a sequence converging to x_0. Hence the limit of the upper bounds, which coincides with the limit of $f(x_n)$, is $f(x_0)$. A similar argument applies to the lower bounds.

A function is said to be continuous throughout an interval if it is continuous at every point of that interval. In order to investigate the properties of a continuous function it is convenient to prove the following lemma, from the theory of sets of points.

The Heine-Borel theorem. If to each point x of a closed interval (ab) we make correspond an interval containing that point as internal point, then a finite number of these intervals cover the whole of (a, b).

Let us prove it first for a one-dimensional interval.



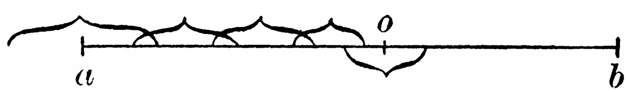
Let d_1 be the interval which corresponds to a and let x_1 be its second endpoint (the first endpoint lies outside (a, b)); let d_2 be the interval corresponding to x_1 and let x_2 be its second endpoint and so on.

Each of the points x_n so obtained has the property that all the points which precede it are internal to a finite number of the given intervals.

We may therefore divide the points of (a, b) into two classes:

(1) the points x such that every interval interior to (a, x) can be covered by a finite number of our intervals;

(2) the points not possessing this property. We shall prove that the points of this second class do not exist.



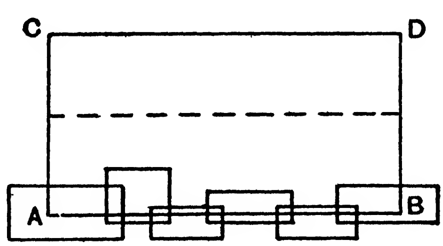
Let O be the last point of the first class; this point clearly exists since if all the points preceding O are of the class (1) so is O.

Now all the points internal to the interval which corresponds to O clearly belong to the first class.

Hence unless O is the last point of (a, b), O cannot have been the last point of the first class.

This proves the theorem.

Let us now prove it for two dimensions. Let $ABCD$ be the given rectangle. By the one-dimensional theorem all the points of AB may be covered by a



finite number of the intervals; let h_1 be the smallest distance from AB of the horizontal sides of these intervals. Then all the points whose height above (a, b) is less than h_1 are internal to a finite number of our intervals.

Let h be the largest number such that all the points of height less than h above AB and which belong to our rectangle are internal to a finite number of our intervals.

The argument of the one-dimensional theorem proves that h is not less than the height above AB of CD.

This proves the theorem in two dimensions and similarly in n.

Uniform continuity. If $f(x)$ is continuous in a closed interval of definition, then given any positive number e, there exists a positive number d, such that if x_1 and x_2 are any two points of the interval of definition distant less than d, then
$$|f(x_1) - f(x_2)| < e.$$

For by the theorem just established, if to every point x_0 we make correspond an interval i surrounding it, such that throughout the concentric interval I twice as large in every direction, we have
$$|f(x) - f(x_0)| < e/2,$$
a finite number of the intervals i cover the whole interval of definition.

Let d be half of the smallest side of these intervals. Then if x_1 is any point of E, x_1 is internal to at least one interval, i_2 say, of the finite set of intervals. The interval I_1 twice as large as i_1 therefore contains all the points distant less than d from x_1, in particular x_2.

If x_0 is the point to which I_1 corresponds, then by adding the inequalities got by writing in turn x_1 and x_2 for x, the theorem follows.

Closely allied to the idea of continuity of a function is that of *uniform convergence of a sequence of functions**.

Let $f_1, f_2, \dots, f_n, \dots$ be any sequence of functions of x†, and let f be their limiting function.

We suppose all these functions bounded and defined in the same interval.

Let us imagine that, besides the coordinates which represent the point x, we have a further coordinate p representing the parameter n, for example its inverse $1/n$.

In this new space we define a new function
$$F(x, p) = f_n(x) - f(x),$$
but only at those points where p represents an integral value of n, and let us agree to give it the value 0 elsewhere.

Then the convergence of the given sequence is said to be uniform at a point x_0, if in our new space the upper and the lower limits of the function $F(x, p)$ at the point which corresponds to $x = x_0$ and $n = \infty$ coincide and are equal to 0; in other words, F is continuous at that point.

This may also be expressed by saying that there is *an unique* $\overline{m+1}$*'ple limit* at x_0, where m is the number of coordinates of x.

* It has been better termed by Du Bois-Reymond *continuous convergence* ("stetige Convergenz").

† Not necessarily continuous.

In the e-language this may be stated as follows:

The given sequence is said to converge uniformly at the point x_0 if, given any positive quantity e, we can find an index n_0 and an interval surrounding x_0, such that throughout it and from and after $n = n_0$ we have

$$|f_n(x) - f(x)| < e.$$

The notion of uniform convergence at a point is of great importance in the theory of functions of real variables; moreover, many properties laboriously deduced from the e-definition in a closed interval become obvious by inspection.

When the convergence is uniform at every point of a given interval (open or closed), it is said to be uniform in that interval. The equivalence of this definition with that given in Chapter I when the interval is closed is easily proved with the help of the Heine-Borel theorem; and it may be shown that we can also define uniform convergence in an open interval as uniform convergence in every closed component.

It is obvious that, if we can determine n_0 such that throughout the closed interval (a, b) and for all n after n_0 we have

$$|f_n(x) - f(x)| < e,$$

then the convergence is uniform at every point of (a, b).

Conversely, if the convergence is uniform at every point of (a, b) then, given e, we can at each point find an n_0 and an interval containing the point, throughout which from and after n_0, we have the above inequality. A finite number of these intervals cover the whole of (a, b); if n_1 is the greatest of the corresponding n_0's then throughout (a, b) from and after n_1 the above inequality holds. Q.E.D.

Theorem. A function $f(x)$ which is the limit of a uniformly convergent sequence of functions of a certain type is also the limit of a monotone ascending and of a monotone descending sequence of functions of that type.

It is sufficient to prove it for the case in which the functions are all defined in a closed interval (a, b).

Let $e_1, e_2, \ldots, e_n, \ldots$ be a monotone sequence of positive numbers tending to zero.

Let f_1 be the first of our functions which throughout (a, b) fulfils the inequality

$$|f_1(x) - f(x)| < e_1.$$

Similarly, let f_n be the first to satisfy the inequality

$$|f_n(x) - f(x)| < e_n$$

throughout (a, b).

The functions $f_n + e_n$ and $f_n - e_n$ both belong to the same type as f_n and the corresponding sequences are respectively monotone ascending and monotone descending and they tend to $f(x)$. This proves the theorem.

Corollary. $f(x)$ belongs to the same type.

If the generating functions are continuous it follows that so is the limiting function. The connection between continuity and uniform convergence is also illustrated by the following theorem.

THEOREM. If a *monotone* sequence of continuous functions has as limit a continuous function, it converges uniformly.

Again take the interval to be closed. Let f_n be the generic term of the sequence and f the limiting function. Suppose the sequence monotone descending. The differences

$$f_n - f$$

form a monotone descending sequence of continuous functions, which are particular cases of L-functions.

By the theorem of bounds the limit of their upper bounds is zero. We can therefore find an n_0, from and after which, it is less than e. Therefore, a fortiori, throughout (a, b)

$$f_n - f < \epsilon \text{ for } n > n_0.$$

Similarly if the sequence is monotone ascending.

2. *The generation of sets of points, and the theory of content.*

A set of points is said to be well-defined if we are given a law which enables us to say of any arbitrarily chosen point whether it belongs to the set or not. If we are given such a set E, we are at once able to define the function which is 1 at every point of the set and 0 elsewhere. Conversely, given any function which only assumes the two values 0 and 1, then the set of points where this function is 1, is well defined. This enables us to generate all well-defined sets of points.

The simplest functions constant in stretches which are not actual constants are those which only assume two values, say, 0 and 1. If such a function is a simple U, the set of points where it is 1 consists of a finite number of closed intervals. We shall call it a simple U-set. The complementary set is a simple L-set.

If we are given a monotone descending sequence of simple U-functions, all of which assume only the values 0 and 1,

$$E_1(x),\ E_2(x),\ \ldots\ E_n(x),\ \ldots,$$

then the set of points E_n, where $E_n(x)$ is unity, is clearly contained in E_{n-1}. We have thus got what we may call *a monotone descending sequence of sets of points*,

$$E_1 > E_2 > \ldots \quad \text{(The symbol $>$ now means containing.)}$$

The set of points E common to all of them will be called *their limiting set* of points, and is identical with the set of points where the function $E(x)$, limit of $E_n(x)$, is unity. Since $E(x)$ is a U-function which assumes only the values 0 and 1, the set of points E will be called a (general) U-set. The complementary set, limit of monotone ascending sequence of simple L-sets (the limiting set now consists of all the points belonging to at least one of the generating sets) is, similarly, a (general) L-set. In this way, proceeding stage by stage as in the generation of functions, we obtain the most general kind of set of points mathematically definable.

Given any set of points E, and a monotone increasing function $g(x)$, the integral of the function $E(x)$ which is 1 in E and 0 elsewhere, may be called

the *total variation* of $g(x)$ with respect to E, or, as we shall say for brevity, *the* $g(x)$*-content* of E.

In order to evaluate the $g(x)$ content of E, we have to find the upper bound of the integrals of the U-functions less than $E(x)$. It is easily seen that it is sufficient to take into consideration those U-functions only, which assume the values 0 and 1 and no others. It is evident that we need only consider non-negative U-functions. Let $a(x)$ be such a function; where $a(x)$ is greater than or equal to $1/n$, we put it equal to 1. This does not affect its upper semicontinuity, nor its property of being nowhere greater than $E(x)$. If we put it equal to 0 where it is less than $1/n$, and make n tend to infinity, we obtain a monotone sequence of U-functions assuming only the values 0 and 1, whose integrals tend to a value greater than the integral of $a(x)$.

It follows immediately that *the content of* E *is the upper bound of the content of* U*-sets contained in* E.

We remark that *a* U*-set is identical with a closed set.* A U-set is closed; this in fact is almost obvious. Let $E(x)$ be the function which is 1 in E and zero elsewhere. Then at a limiting point of E, the function $E(x)$ cannot assume a value less than 1 since it is upper semicontinuous. Conversely, every closed set is a U-set*. For let E be the closed set and let $E(x)$ be 1 in E, 0 outside. Then $E(x)$ is upper semicontinuous. For a point where $E(x)$ is zero is internal to an interval containing no point of E (if such an interval did not exist the point would be a limiting point of E, i.e. $E(x)$ would be 1); therefore at such a point, the upper and lower limits of $E(x)$ are both zero. On the other hand, at a point where $E(x)$ is 1, $E(x)$ is obviously not less than the upper and lower limits. Hence $E(x)$ is upper semicontinuous and E is a U-set.

To denote the $g(x)$ content of E, we use the symbol $m_g(E)$. m stands for measure, which some writers use instead of content.

Historically, although, as we have seen, the general idea of the theory of integration follows naturally from the ideas of Cauchy and Darboux, the theory of content preceded the modern theory of integration. This was the main reason why mathematicians of the old school did not at once recognise its value.

We now proceed to prove some important theorems.

THEOREM. If $f(x)$ is zero except at a set of points of g-content zero, the integral of f with respect to g vanishes.

Let E be the set of points where $f(x)$ is not zero, and $E(x)$ the function which is 1 throughout that set and zero elsewhere. Then, clearly, $f(x)$ equals $f(x)\times E(x)$. Therefore

$$\int f(x)\,dg=\int f(x)\,E(x)\,dg.$$

But†

$$\int f(x)\,E(x)\,dg=\int f(x)\,d\left\{\int E(x)\,dg\right\},$$

* This is equivalent to the following theorem of Cantor: A closed set consists of the points non-internal to a set of non-overlapping intervals.

† By Theorem I, p. 40.

or, since the integral of $E(x)$ is zero in every interval,

$$\int f dg = 0. \qquad \text{Q.E.D.}$$

THEOREM. If

$$\int f dg = 0$$

in every interval, then $f(x)$ is zero except, at most, at a set of g-content zero.

Since $\int f dg$ is the uniform limit of the integral of the function f_n identically equal to zero, and since f_n, being constantly zero, remains less than an integrable function, it follows from Theorem IV of p. 37 that the integral of $|f|$ is the uniform limit of that of $|f_n|$, that is, zero. We may therefore suppose f to be non-negative. Now, the set of points E_n where f is not less than $1/n$ has content zero because otherwise

$$0 = \int f dg \geqslant \frac{1}{n} \int E_n(x)\, dg > 0$$

by the fundamental property of integration III.

Now the set of points E where f is not zero is the limit of E_n as n tends to infinity, for it consists of all the points belonging to at least one E_n. By the theorem on term by term integration of a monotone sequence, the g-content of E is zero.

3. *Functions defined in a set of points.*

So far we have only considered functions which are defined at every point of an interval. More generally, we may suppose the values of a function to be given at some points, unknown at others. Such a function is *defined in a set of points.*

y is said to be function of x in a set of points E, if to each x belonging to that set there corresponds a single value of y.

A function which is defined in an interval is obviously also defined in every set of points contained in that interval.

In speaking of a function defined in a set of points E we merely exclude all consideration of other points, whether or not the function is defined at any of the latter.

A function defined in a set of points E has for upper bound in an interval D, the upper bound of those of its values which correspond to points of E inside the interval D.

With this convention, we may also define upper and lower limits of our function at every point x_0 such that in each interval containing x_0 there is a point of E (other than x_0): on the other hand, if there is an interval containing x_0 but no other point of E, then it is clearly impossible to define the upper and lower bounds of our function in that interval excluding x_0, and consequently impossible to define upper and lower limits at x_0 in this case. Hence the upper and lower limits of a function defined in a set of points E exist solely at points of the set E' of the limiting points of E; at any such point x_0 (which need not of course belong to E) they are defined to be the unique

limits of the upper and lower bounds, respectively, of f in any succession of intervals each containing the point x_0 as internal point and shrinking up to x_0 (the point x_0 being, as before, excluded in the evaluation of the bounds in each interval).

Upper and lower semicontinuity. Continuity. A function is said to be upper semicontinuous at a point x_0 where it is defined, provided its upper limit at that point (if existent) is not greater than its value at that point.

A function is said to be continuous at a point if it is both upper and lower semicontinuous at that point.

Thus continuity at a point depends on the set of points at which the function is defined. A function, defined in E, may be discontinuous at a point x_0 of E and yet become continuous at x_0 when considered to be defined in a subset E_1 of E.

A function is said to be *continuous over a set of points* E, if, when considered as defined only in that set E, it is continuous at every point of E.

A function is said, *at a point* x_0, to be *continuous with respect to a set of points* E, if, when considered as defined only in that set E, it is continuous at x_0. Obviously a function which is continuous over a set of points E is continuous over every component of E.

This may also be expressed in the e-language: f is continuous over E if, to every point x_0 of E, we can make correspond an interval surrounding x_0, throughout which, for all points x of E, we have

$$|f(x)-f(x_0)|<e.$$

The Generalised Heine-Borel Theorem. If to each point x of a closed set of points E we make correspond an interval containing that point as internal point, then a finite number of these intervals cover the whole set E.

Let (a, b) be an interval containing E, and let E' be those of its points which do not belong to E. Let (a, b) be closed. To every point of E' we can make correspond an interval surrounding it containing no points of E.

To each point of (a, b) now corresponds an interval. A finite number of these intervals cover the whole of (a, b). Of these intervals those corresponding to points of E' contain no points of E; the others cover the whole of E.

Q.E.D.

Uniform continuity. If $f(x)$ is continuous over a closed set of points E, then, given any positive number e, there exists a positive number d, such that, if x_1 and x_2 are any two points of E distant less than d, then

$$|f(x_1)-f(x_2)|<e.$$

The proof is but a repetition of the one on p. 44 *mutatis mutandis*.

THEOREM. If $f(x)$ is continuous over a *closed* set of points E, it assumes its upper and its lower bound in that set.

The argument is exactly the same as for the corresponding theorem for a semicontinuous function in an interval. The point constructed, being the limit of a set of intervals all of which contain points of E, certainly belongs to E since E is closed.

COROLLARY. An interval cannot be the sum of two closed sets of points without common points*.

The distance between two points is a continuous function of position and assumes its lower bound as the two points vary in the two sets. This lower bound cannot therefore be zero.

THEOREM. The set of points of E at which $f(x)$ assumes values not less (or not greater) than K is closed.

At a limiting point of such points, $f(x)$ by continuity assumes a value not less (not greater) than K.

COROLLARY. A function which is continuous in a closed interval† there assumes all values between its upper and its lower bound. For let K be any such a value, then the two closed sets where it is $\geqslant$ and where it is $\leqslant K$ must by the corollary above have common points.

Successions of functions.

Just as we considered in Chapter I successions of functions defined in an interval, we may consider successions of functions defined in a set of points E, and apply to them the investigation of Chapter I.

In the case when the set of points is closed, the convergence of the functions $f_n(x)$ in that set of points is said to be *uniform*, if, given any positive number e, an index N can be found, independent of x, such that, for every n from and after N,

$$|f_n(x)-f(x)|\leqslant e,$$

for every x belonging to the set of points.

More generally we may define uniform convergence as in Appendix, 1.

THEOREM. A sequence of functions, continuous over a closed set E, and which converges uniformly over E, converges to a function which is also continuous over E.

Proof:

$$|f(x_1)-f(x_2)|\leqslant|f_N(x_1)-f(x_1)|+|f_N(x_2)-f(x_2)|+|f_N(x_1)-f_N(x_2)|.$$

Choose N so that $|f_N(x)-f(x)|\leqslant e/3$ for all the x's of E, which we can, because of the uniform convergence, and δ so that $|f_N(x_1)-f_N(x_2)$ may be less than $e/3$ as soon as x_1 and x_2 are distant less than δ and belong to E; then

$$|f(x_1)-f(x_2)|\leqslant e.$$

The generation of functions defined in a set of points.

Here proceed as in Chapter II. The simplest functions defined in E are those which are constant in E. The next simplest are the functions constant in stretches, i.e. if (a, b) be any interval containing E, it is the sum of a finite number of stretches, inside each of which at all the points belonging to E the function is a finite constant.

* A set of points which is not expressible as the sum of two closed components without common points is called a continuous set of points (Jordan).

† More generally "over a continuous set of points."

Then as in Chapter II we generate by monotone sequences the most general kind of function it is possible to define in E. We can also shew that, if E is closed, a uniformly convergent sequence of functions of a certain type generates a function of the same type*.

We now come to a remarkable theorem (Egoroff).

A function defined in a set of points E and obtainable by monotone sequences, if finite throughout that set, is continuous over a closed subset C_n, whose g-content differs from that of E by less than $1/n$, n being an arbitrary positive integer.

This theorem is an immediate corollary to the following theorem:

THEOREM. A sequence of functions, defined in a set of points E, and having a finite limit, will converge uniformly in a closed set of points C_n whose g-content differs from that of E by less than $1/n$.

Let $f_n \to f$ in a set of points E where f assumes only finite values, and let e and d be any pair of positive numbers.

Let E_r denote the set of points at which

$$|f_{r+h} - f| < e$$

for all positive integral values of h.

Obviously, $$E_1 < E_2 < \dots;$$

also, every point of E belongs to some E_r, since the succession f_n converges in E and f assumes only finite values in E. Therefore

$$E_1 < E_2 < \dots \to E.$$

Hence by the theorem on term by term integration

$$m_g(E_n) \to m_g E.$$

Hence since $m_g(E)$ is finite we can determine N so that

$$m_g(E_N) > m_g(E) - d.$$

Then in E_N $$|f_{N+h} - f| < e$$

for all h.

Let us choose for e a value $1/p$, p being a positive integer, and let us choose for d a value $\epsilon/2p$, ϵ being a positive number independent of p.

Let the value of N obtained be denoted by N_p, supposing ϵ fixed for the present.

Let $\mathcal{E}_1$ be the set E_{N_1},

$\mathcal{E}_2$ the set common to E_{N_1}, E_{N_2}, which may be called their product $E_{N_1} . E_{N_2}$,

$\mathcal{E}_3$ the set common to E_{N_1}, E_{N_2}, E_{N_3},

and so on.

Clearly $$\mathcal{E}_1 > \mathcal{E}_2 > \mathcal{E}_3 \dots \to \mathcal{E} \text{ say.}$$

Also $$m_g(E - \mathcal{E}_1) < \epsilon/2 \qquad < \epsilon$$

$$m_g(E - \mathcal{E}_2) \leqslant m_g(E - \mathcal{E}_1) + m_g(E - E_{N_2}) \leqslant \epsilon\left(\tfrac{1}{2} + \tfrac{1}{4}\right) \qquad < \epsilon$$

$$m_g(E - \mathcal{E}_{n+1}) \leqslant m_g(E - \mathcal{E}_n) + m_g(E - E_{N_{n+1}}) < \epsilon\left(\frac{1}{2} + \frac{1}{4} + \dots + \frac{1}{2^{n+1}}\right) <$$

* This of course contains the theorem of the preceding page as a particular case.

But by the theorem on term by term integration

$$m_g(\mathcal{E}_n) \to m_g(\mathcal{E}).$$

But $\mathcal{E}$ belongs to E_{N_p} whatever p. In $\mathcal{E}$, therefore,

$$|f_{N_p+h} - f| < \frac{1}{p},$$

whatever p.

Hence the sequence f_n converges uniformly on a set of points $\mathcal{E}$ whose g-content differs from that of E by less than ϵ.

Now the content of $\mathcal{E}$ is the upper bound of that of its closed components. Hence there is a closed set C whose g-content differs from that of E by less than 2ϵ, in which the sequence f_n converges uniformly to its limiting function.

Q.E.D.

INDEX

Absolutely convergent integral 32
analytical 14, 15
area v, 22

Baire vii
Borel vii; Heine-Borel theorem 43, 49
bounds 1, 5; theorem of 8

Calculus v, vi
Cantor vi, vii; Cantor-Dedekind axiom 3
Cauchy v, vi, 24, 42, 47
constant 10
content vii, 46 seq.
continuity vi, 7, 42, 49; continuous convergence, see uniform convergence; continuous set 50
convergence 2, 3, 5

Darboux v, 24 seq.
Dedekind 3
differentiation vi
discontinuous vi, 7
Du Bois-Reymond 44

e-definitions 3, 7, 42, 44
Egoroff 51
evaluation 24
exhaustion v

Finite 1; set 4
function v, vi, 4; simple 9; upper, lower 4; L-, U-functions 7, 11; LU-, UL-, etc. 13; functions mathematically defineable 15

Galois vi
generation of functions 9 seq., 50; of sets of points 47 seq.
Grassmann vi
Greeks v, 3, 22

Heine 42, 43; Heine-Borel 43, 49

Increment 19
infinitesimal v, vi
integrable 32
integral, see integration
integrand 20
integration, see table of contents; properties of 20, 21, 28, 32
integrator 20, 40 seq.
interval 4
isolated 4

Jordan 50
jump 6

Lebesgue vii
Leibnitz v
L-function 7; LU, LUL, etc. 12 seq.
limit v, 1, 2, 5, 6; limiting point 3; limiting function 5

Maximum limit 6
measure, see content
minimum limit 6
monotone, see table of contents

Newton v
non-differentiable vi

Properties of class, fundamental 10 seq.; of integration 21, 28, 32
pseudo-monotone 17, 18

Quadrant 6

Related 25
repeated 35
Riemann v, vi, 24

Semi-continuous 7, 49
sequence 2
series v
set vii, 1, 4, 46 seq.
simple L- or U-function 9
Stieltjes 23
stretch 4; constant in 9
subsequence 2
subsuccession 2
succession, see table of contents

U-function 7; UL, etc. 13 seq.
uniform continuity 43; convergence 5, 44

Weierstrass v, vi

Young vii

Zermelo 42